



Quelques contributions à l'imagerie médicale multimodale et interventionnelle

Erwan Kerrien

► To cite this version:

Erwan Kerrien. Quelques contributions à l'imagerie médicale multimodale et interventionnelle. Imagerie médicale. Université de Lorraine, 2018. tel-01754699

HAL Id: tel-01754699

<https://inria.hal.science/tel-01754699>

Submitted on 30 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quelques contributions à l'imagerie médicale multimodale et interventionnelle

THÈSE

présentée et soutenue publiquement le 9 janvier 2018

pour l'obtention d'une

Habilitation de l'Université de Lorraine
(mention informatique)

par

Erwan KERRIEN

Composition du jury

Président : Antoine TABBONE, Professeur à l'Université de Lorraine

Rapporteurs : Isabelle BLOCH, Professeur à Télécom ParisTech
Jocelyne TROCCAZ, Directrice de Recherche CNRS
Pascal HAIGRON, Professeur à l'Université de Rennes I

Examineurs : Marie-Odile BERGER, Directrice de Recherche Inria (garante scientifique)
Antoine TABBONE, Professeur à l'Université de Lorraine
Stéphane COTIN, Directeur de Recherche Inria

Table des matières

1	Introduction	7
1.1	L'imagerie médicale multimodale et interventionnelle	7
1.2	Contributions scientifiques	8
1.2.1	Radiologie interventionnelle	8
1.2.2	La coeliochirurgie	10
1.2.3	L'inversion acoustico-articulatoire	11
1.2.4	Aspects méthodologiques	12
1.3	Plan du mémoire et Guide de lecture	13
2	Acquisition et traitement de données articulatoires	15
2.1	Introduction	15
2.1.1	Contexte scientifique	15
2.1.2	Contexte applicatif	17
2.1.3	Contributions	19
2.2	Le système ARTIS : description et principe	19
2.2.1	Présentation générale	19
2.2.2	Echographie	20
2.2.3	Capteurs électromagnétiques	20
2.2.4	Acquisition stéréovision	22
2.2.5	PC de contrôle	22
2.3	Recalage et synchronisation des données EM et US	22
2.3.1	Données et repères	23
2.3.2	Utilisation d'un fantôme	24
2.3.3	Procédure de calibration avec notre fantôme	25
2.3.4	Synchronisation EM/US/audio	28
2.4	Acquisition et recalage des données IRM	30
2.4.1	Contexte bibliographique	30
2.4.2	Protocole IRM proposé	31
2.4.3	Recalage des images IRM	34
2.5	Quelle fiabilité pour ces données ?	35
2.5.1	Évaluation visuelle	35
2.5.2	Récapitulatif des repères et transformations	36
2.5.3	Évaluation quantitative	37
2.6	Exemple d'exploitation du système ARTIS	38
2.7	Recalage et segmentation automatique d'images IRM	39
2.7.1	Hypothèses et données	39
2.7.2	Méthode	41
2.7.3	Résultats	43
2.8	Conclusion et Perspectives	43

3	Vers une navigation augmentée en neuroradiologie interventionnelle	47
3.1	Introduction	47
3.1.1	Contexte	47
3.1.2	Contexte médical	48
3.1.3	Contributions	49
3.2	Évaluation clinique de l'imagerie péri-opératoire 3DDSA	51
3.2.1	Premier contexte clinique : traitement endovasculaire des anévrismes intracrâniens	52
3.2.2	Second contexte clinique : traitement des MAV par radiochirurgie	54
3.3	Modélisation et calibration de l'arceau vasculaire	59
3.3.1	Contexte	59
3.3.2	Données capteurs et géométrie idéale de l'arceau	59
3.3.3	Étude et modélisation de l'arceau	63
3.4	Fluoroscopie Augmentée : concept et évaluation	66
3.4.1	Rappel du contexte clinique	67
3.4.2	Principe et prototype de FA	68
3.4.3	Protocole de validation du prototype de FA	68
3.4.4	Résultats et discussion	69
3.5	Reconstruction du micro-guide par stéréovision	71
3.5.1	Contexte	71
3.5.2	Reconstruction stéréo de courbes	71
3.5.3	Principe de l'algorithme proposé	73
3.5.4	Pré-sélection des fragments 3D	74
3.5.5	Ordonnancement des fragments 3D en supports candidats	76
3.5.6	Résultats et discussion	77
3.6	Conclusion et perspectives	79
4	Contributions à la simulation interactive pour la neuroradiologie interventionnelle	81
4.1	Introduction	81
4.1.1	Contexte	81
4.1.2	Contributions	84
4.2	Segmentation vasculaire : l'algorithme RBT	85
4.2.1	Description générale de l'algorithme	85
4.2.2	Extraction des points candidats	85
4.2.3	Ajustement de cylindre par RANSAC	86
4.2.4	Construction de l'arbre	87
4.2.5	Validation	89
4.3	Reconstruction vasculaire : l'algorithme LIM	93
4.3.1	Modèle implicite	93
4.3.2	Formulation énergétique	96
4.3.3	Optimisation	97
4.3.4	Validation	99
4.4	Application à la simulation : implantation et validation	101
4.4.1	Contexte de la simulation	102
4.4.2	Implantation de notre modèle dans un contexte simulation	102
4.4.3	Quelques éléments pour la validation d'une simulation interactive	107
4.5	Conclusion et perspectives	110
5	Conclusions et perspectives	113
5.1	Conclusion générale	113
5.2	Vers une simulation guidée par l'image	113
5.2.1	Détecter	114
5.2.2	Prédire	115

5.2.3 Adapter	116
Bibliographie générale	117
Publications de l'auteur	134
Chapitres d'ouvrage	134
Actes de conférences	134
Revues internationales	134
Conférences internationales	135
Conférences nationales	137
Communications	138
Conférences médicales	138
Brevets	139
Rapports	140
Logiciels	140
Thèses et mémoires	140
Vulgarisation scientifique	140

Chapitre 1

Introduction

Sommaire

1.1	L'imagerie médicale multimodale et interventionnelle	7
1.2	Contributions scientifiques	8
1.2.1	Radiologie interventionnelle	8
1.2.2	La coelochirurgie	10
1.2.3	L'inversion acoustico-articulatoire	11
1.2.4	Aspects méthodologiques	12
1.3	Plan du mémoire et Guide de lecture	13

1.1 L'imagerie médicale multimodale et interventionnelle

L'imagerie médicale, expression que j'emploierai dans ce manuscrit pour désigner sa déclinaison dans les sciences informatiques, repose sur des méthodologies que l'on retrouve traditionnellement en vision artificielle : segmentation, recalage, suivi et détection de changements, reconstruction géométrique et ajustement de modèles, mesure de paramètres à partir des images. Mais trois éléments spécifiques interdisent de considérer l'imagerie médicale comme un simple domaine d'application de la vision artificielle. La diversité des phénomènes physiques exploités pour produire des images justifie de s'intéresser aux particularités des processus physiques sous-jacents (artéfacts, bruit, ...) et d'avoir une approche multimodale. La diversité des domaines applicatifs médicaux implique une variété équivalente de systèmes, de contextes, de protocoles, d'objectifs, et rend nécessaire une approche inter-disciplinaire. Enfin, la complexité des structures imagées : ces structures anatomiques, en général multiples et hétérogènes, possèdent une forme qui est difficilement capturée par les descripteurs utilisés habituellement en analyse d'image (détecteurs de coins, segments, descripteurs de texture) [MIH11].

L'imagerie médicale est de plus un de ces rares domaines où nous disposons d'emblée d'acquisitions volumétriques. Le courant des années 1990 a en effet été le témoin d'une révolution provoquée d'une part par les progrès en reconstruction d'images et notamment l'avènement de l'imagerie par résonance magnétique (IRM), et par la démultiplication des moyens de calcul informatiques d'autre part [DA00]. La puissance informatique devenue accessible dans ces années a rendu envisageable de traiter, en un temps compatible avec la clinique, ces données massives et multimodales. Car outre l'IRM, les médecins avaient à leur disposition d'autres modalités d'imagerie 3D plus anciennes, comme la tomodensitométrie par rayons X (TDM, connue par l'acronyme anglais CT pour *Computed Tomography*), ou les imageries nucléaires fonctionnelles (notamment la tomographie par émission de positons (TEP)), ainsi que d'autres, en devenir comme l'échographie 3D et l'angiographie 3D (3DRA pour *3D Rotational Angiography*) [10].

À la fin de ma thèse, en 2000, les mécanismes de transfert des avancées en vision artificielle vers l'imagerie médicale étaient bien en place et par ailleurs ce domaine possédait déjà un premier corpus de résultats spécifiques. Un excellent exemple est le problème de recalage qui est une clé essentielle permettant la multimodalité. Le contexte médical a suscité des approches

originales pour répondre à cette question [CMD⁺95, VWI95]. On pouvait alors la considérer pour l'essentiel résolue dans sa version rigide, statique, et pour des modalités de nature similaire, c'est-à-dire des images 3D et/ou 2D, que ce soit par la mise en correspondance d'indices structuraux extraits des images ou par l'optimisation d'un critère de similarité entre les images [MV98]. En revanche, des avancées étaient encore nécessaires dans un contexte non-rigide [Hol08] ou interventionnel [MTLP12].

La chirurgie connaît en effet depuis quelques années une véritable révolution. Ce terme, en apparence fort, se justifie par l'hiatus entre la pratique actuelle de la chirurgie et l'image qu'en a le grand public. En effet, une grande partie des opérations se réalisent aujourd'hui de manière minimalement invasive, et cette proportion croît inéluctablement. L'expression «minimalement invasif» recouvre un principe simple : réparer, traiter, opérer à l'intérieur du patient, depuis l'extérieur, à travers une ouverture la plus petite possible. La conséquence pour le patient est largement bénéfique : traumatisme moindre, récupération plus rapide, temps d'hospitalisation plus court, risque réduit d'infection nosocomiale, etc. En revanche, le chirurgien doit de son côté faire face à une dégradation importante des conditions d'opération : plus haute technicité du geste chirurgical, champ de vision réduit, pas de vue tridimensionnelle, pas de retour tactile (palpation impossible), entre autres.

Ces techniques regroupent deux grandes familles de pratiques chirurgicales et thérapeutiques récentes : la radiologie interventionnelle et la chirurgie endoscopique. Leur développement dans la clinique date de ces trente dernières années, mais ce n'est qu'à la fin des années 1990 qu'elles ont atteint une maturité telle que la communauté de recherche en imagerie médicale informatique a vu dans l'imagerie interventionnelle un de ses domaines applicatifs majeurs [DA00]. Cette maturité a reposé sur deux critères : l'un technologique, avec la stabilisation des technologies d'imagerie (caméras endoscopiques, et rayons X faible dose (fluoroscopie numérique)) et le développement des outils, endoscopiques et interventionnels, et l'autre médical, avec la reconnaissance de l'impact positif de ces pratiques d'un point de vue clinique et donc leur large diffusion dans les centres hospitaliers. Ce contexte est reconnu aujourd'hui offrir des challenges spécifiques et nombreux au domaine de la vision par ordinateur [MIH11].

L'objectif de mes recherches est de rendre au praticien un environnement de travail intuitif où il peut à nouveau se concentrer sur l'aspect médical du geste et ne pas voir sa réflexion polluée par des questions purement techniques. L'objectif final est de proposer un environnement où le chirurgien ou le radiologue interventionnel pourra à nouveau concevoir, planifier, répéter et réaliser son geste en 3D. Cela requiert d'aborder différentes méthodologies scientifiques, mais demande aussi d'intégrer le contexte applicatif à toutes les étapes de la recherche sous peine d'être totalement hors-sujet du point de vue du médecin.

1.2 Contributions scientifiques

Située dans le contexte de l'imagerie médicale interventionnelle et multimodale, mon activité de recherche a toujours été guidée par l'application, et s'enracine dans une approche systématiquement pluridisciplinaire. Les questions qui la suscitent n'émergent pas d'un substrat théorique fixé, dont les avancées pourraient bénéficier à une application parmi d'autres, mais surgissent à l'interface entre disciplines. Une part importante de mon activité consiste donc en un travail d'assimilation du cadre applicatif, afin d'identifier et d'exprimer ses contraintes effectives qui fonderont le caractère scientifique du problème.

Je donne donc ici un résumé de mes activités de recherche suivant les applications qui les ont suscitées, en terminant par une synthèse méthodologique.

1.2.1 Radiologie interventionnelle

La radiologie interventionnelle par cathétérisme est une technique de chirurgie dite minimalement invasive : une simple incision est faite au niveau de l'aîne, par laquelle un *cathéter* (long tube fin et flexible) est inséré dans l'artère fémorale, puis est guidé à travers le réseau vasculaire jusqu'à la pathologie. Le cathéter sert alors de canal pour acheminer les micro-outils qui vont

permettre de réaliser le traitement. La neuroradiologie interventionnelle en est une sous-spécialité qui a fourni un cadre applicatif particulier et privilégié à mes recherches. Elle vise les pathologies vasculaires du système nerveux central (crâne, cou, rachis). Les artères y sont donc peu, voire pas déformables, et les difficultés se situent à deux niveaux : d’une part les structures vasculaires et les pathologies ont souvent une forme très complexe en trois dimensions, et d’autre part la grande précision requise (sub-millimétrique) et la complexité du comportement mécanique du micro-outil rendent le geste médical très délicat.

Comme pour tout geste chirurgical, la manipulation des outils interventionnels doit se concevoir en trois dimensions (3D). Cependant, le médecin contrôle son geste grâce à de la radiographie temps réel à faible dose de rayons X appelée *fluoroscopie*. La fluoroscopie n’offre néanmoins qu’une piètre qualité d’image, en deux dimensions (2D), alors que le médecin dispose de plusieurs imageries 3D pour analyser la pathologie. En particulier, l’angiographie rotationnelle 3D (3DRA) peut offrir des instantanés 3D du site opératoire, mais pas de manière continue. La question fondamentale qui soutient cet axe de recherche depuis ma thèse [71] est donc celle de la fusion de ces modalités (auxquelles s’ajoutent éventuellement l’IRM) afin que le médecin bénéficie à la fois des trois dimensions de la 3DRA et de la dynamique de la fluoroscopie.

Cette recherche a débuté en 1996 dans une approche volontairement fortement pluridisciplinaire. Une étroite et active collaboration s’est ainsi établie avec les neuroradiologues interventionnels du Département de Neuroradiologie Interventionnelle Diagnostique et Thérapeutique du CHU de Nancy, sous l’impulsion initiale du Prof. Picard et l’action sans faille du Prof. Anxionnat, ainsi qu’avec la société GE Healthcare, fabriquant du matériel d’imagerie. Depuis 2007, nous collaborons également sur ce sujet avec l’équipe Inria SHACRA, qui a depuis donné naissance à l’équipe DEFROST à Lille, dirigée par Christian Duriez, et l’équipe MIMESIS de Strasbourg, dirigée par Stéphane Cotin et dont je suis membre associé.

Collaboration avec le CHU de Nancy. Imagerie entièrement nouvelle en 1997, une étape cruciale a été de valider la 3DRA cliniquement. Nous l’avons fait pour le traitement des deux principales pathologies vasculaires cérébrales, à savoir l’embolisation des anévrysmes intracrâniens par coils [10, 56] et la radiochirurgie des malformations artério-veineuses [Anx03, 8, 35, 36, 37, 54, 55], sujet au centre de la thèse du Prof. René Anxionnat [Anx03], (soutenue en 2003, dirigée par Marie-Odile Berger et que j’ai co-encadrée). Les contributions de ces travaux sont des méthodologies de validation perceptive, quantitative et statistique d’une modalité d’imagerie dans un but clinique défini. Au-delà des contributions scientifiques, cette collaboration permet de bâtir une compréhension et une confiance réciproques essentielles à des avancées utiles aux médecins.

Collaboration avec le CHU de Nancy et GE Healthcare. J’ai développé pendant ma thèse [71] (contrat Cifre avec GE Healthcare) des algorithmes de recalage de la 3DRA avec des volumes d’IRM (problème 3D/3D) [38] ainsi qu’avec des images d’angiographie (problème 3D/2D) [39, 40]. Afin d’initialiser le recalage 3D/2D de manière fiable et précise, j’ai par ailleurs modélisé le comportement mécanique de l’arceau vasculaire par des méthodes de vision par ordinateur et montré qu’une précision de quelques millimètres pouvait être atteinte [41]. Ce travail a donné lieu à 6 brevets [61, 62, 63, 64, 65, 66], et un produit de fusion 3DRA/IRM, commercialisé par GE Healthcare en 2001 alors que j’y travaillais comme ingénieur de recherche.

Nous avons actualisé ce modèle suite à l’arrivée des panneaux plats (capteurs CCD pour rayons X), grâce à une méthode de calibration inspirée de l’ajustement de faisceaux [33, 51]. L’étape suivante consistait à tirer profit de ce recalage pour offrir au médecin une vue fusionnée, disponible en salle d’opération pendant le traitement. Notre application du concept de Réalité Augmentée au contexte de la neuroradiologie interventionnelle a donné naissance à la Fluoroscopie Augmentée. Nous avons montré l’intérêt de cette fusion à travers une étude médicale préliminaire [51]. Ce travail a été réalisé dans le cadre de la thèse de Sébastien Gorges (contrat Cifre avec GE Healthcare, soutenue en 2007, dirigée par Marie-Odile Berger et que j’ai co-encadrée). Il a mené à un produit apparu au catalogue de GE Healthcare en 2008 et a donné

lieu à 4 brevets [57, 58, 59, 60].

La visualisation par Fluoroscopie Augmentée est cependant limitée à l'incidence de vue de la fluoroscopie. Nous nous sommes alors concentrés sur l'information pertinente contenue dans l'image fluoroscopique, le micro-outil, dont une segmentation était disponible grâce à des travaux internes à GE Healthcare. Nous avons proposé une reconstruction visuelle du micro-guide à partir de deux vues, dont on dispose sur un système biplan, sous forme de recherche d'un chemin dans le graphe d'hypothèses d'appariements [15] et étudié des trajectoires à faible nombre de vues pour une reconstruction tomographique d'un coil [14]. Ces travaux ont fait l'objet de la thèse de Charlotte Delmas (contrat Cifre avec GE Healthcare, en cours de finalisation, et que j'ai co-dirigée avec Marie-Odile Berger).

Collaboration avec le CHU de Nancy et SHACRA (devenu MIMESIS et DEFROST).

De la rencontre avec les chercheurs de l'équipe SHACRA en 2007 (Stéphane Cotin, Christian Duriez et Jérémie Dequidt) est née l'idée de faire appel à la simulation physique en temps interactif pour appliquer des contraintes de forme réalistes à la modélisation du micro-outil à partir des images fluoroscopiques. Le premier défi était de réussir à simuler le comportement d'un micro-outil dans les artères, à la fréquence de la fluoroscopie (15 images par seconde), tout en préservant un niveau de réalisme poussé [24]. Mon rôle consistait à modéliser le réseau artériel à partir des données patients [26]. Un tel modèle surfacique définit en effet les conditions aux limites de la simulation. L'arbre vasculaire est cependant d'une forme très complexe. Un modèle précis implique a priori un modèle complexe qui a un impact négatif sur la rapidité des calculs nécessaires à la simulation, notamment pour gérer les collisions. Par exemple, nos premières tentatives utilisant des surfaces actives triangulées [50] ont dû être abandonnées suite aux problèmes numériques constatés lors de la simulation.

Notre approche exploite le fait que la gestion de collisions n'implique que des entités locales : un point et la portion voisine de la surface. Un arbre de fonctions implicites locales correspond à ces caractéristiques : la structure de l'arbre est fournie par la ligne centrale des vaisseaux et une fonction implicite est associée à chaque point de cette ligne centrale. Chaque surface implicite se confond localement avec le vaisseau sanguin et peut servir de succédané local à la surface globale pour la gestion des collisions. Pour extraire les lignes centrales, nous avons développé un nouvel algorithme de suivi de vaisseau, nommé RANSAC-Based Tracking (RBT) [21], basé sur une méthode d'estimation robuste. Dans un deuxième temps, nous avons employé pour fonction implicite locale une somme de sphères implicites (*Blobby Model* [Mur91]). Nous avons amélioré l'existant en proposant l'algorithme de Local Implicit Modeling (LIM) reposant sur une fonction de coût à minimiser par un affinage progressif du modèle [20]. Une validation extensive de nos travaux a été faite, en particulier l'apport de notre modélisation a été prouvé par rapport aux modèles triangulés usuellement employés [5].

Cette approche a notamment été développée dans le cadre de la thèse d'Ahmed Yureidini [Yur14] (soutenue en 2014, dirigée par Stéphane Cotin et que j'ai co-encadrée). Ce travail a été soutenu par les projets SIMPLE (2006-2007, *SIMulation for PLanning the Embolization of intracranial aneurysms*, action de recherche coopérative Inria que j'ai coordonnée), Sofa-InterMeds (2009-2012, action d'envergure nationale Inria de soutien à la plateforme Sofa, pilotée par Stéphane Cotin) et IDEaS (2012-2015, *Image Driven Simulation*, ANR jeune chercheur, piloté par Jérémie Dequidt).

1.2.2 La coeliochirurgie

La coeliochirurgie est une autre approche thérapeutique minimalement invasive. Elle consiste à opérer par le biais de deux trocars insérés dans l'abdomen via de petites incisions cutanées. Le geste se fait grâce au retour visuel d'une caméra vidéo montée sur un troisième trocar. De nouvelles difficultés viennent s'ajouter à celles évoquées plus haut. L'anatomie est en effet très complexe (tous les viscères et leur réseau vasculaire), et fortement déformable. Se posent alors des problèmes de modélisation géométrique et mécanique, et de mise en correspondance entre l'image per-opératoire et le modèle 3D.

Un projet, soutenu par plusieurs stages financés par la région Lorrain, et réalisé en collaboration avec le Service d’Urologie (Prof. Hubert) et le Service d’Imagerie Guilloz (Prof. Blum) du CHU de Nancy, s’est intéressé à la chirurgie du rein. Le problème des médecins était de distinguer les artères, des veines et de l’uretère et de mieux appréhender leurs positions par rapport au rein. Nous avons proposé un algorithme de segmentation de ces structures à partir de 3 scanners acquis à des temps différents lors de l’injection d’un produit de contraste, pour capturer l’opacification successive de chaque structure. Notre méthode recalcule de manière non-rigide ces trois acquisitions avant d’en faire la segmentation successive des artères, veines et uretère [32, 53].

Ces modèles 3D de l’anatomie sont utiles pour planifier l’opération et se familiariser avec le site chirurgical, mais ils pourraient aussi servir à pallier les limitations de la vue coelioscopique et notamment les occultations sévères dont elle souffre. Par exemple, dans le cas du foie, les structures sensibles (réseau vasculaire) ou d’intérêt (tumeur) sont internes à l’organe. La forte déformation que peut subir le foie, et donc ses structures internes, pendant l’opération met à mal une solution de Réalité Augmentée. Nous avons suivi une approche de simulation guidée par l’image, où un modèle biomécanique du foie et de son réseau vasculaire subit une déformation de manière à ce que la surface du modèle géométrique du foie s’adapte à un nuage de points reconstruits à partir de la paire d’images coelioscopiques (l’endoscope est stéréoscopique). L’attraction de la surface par les points se fait par l’intermédiaire de forces de ressort dont la raideur est reliée au hessien au point d’intérêt détecté par un détecteur SURF [7, 18, 19, 22]. Ce travail a été réalisé dans le cadre de la thèse de Nazim Haouchine [Hao15] (soutenue en 2015, co-dirigée par Stéphane Cotin et Marie-Odile Berger).

1.2.3 L’inversion acoustico-articulatoire

Un troisième cadre applicatif abordé est celui de l’inversion acoustico-articulatoire. L’intelligibilité d’un signal de parole ne dépend pas uniquement de la qualité du son. Nous utilisons également en permanence des indices visuels portés par le visage. Le problème de la «tête parlante» pose la question d’animer une tête virtuelle en y reproduisant les indices visuels complémentaires du signal de parole. Les applications en sont variées : jeux, films d’animation, l’usage d’avatars pour des téléconférences à forte audience, ou communications en milieu bruyant. La synthèse d’une tête parlante passe par l’inversion acoustico-articulatoire dont l’objectif est de synthétiser le conduit vocal (des cordes vocales aux lèvres) à partir du seul signal de parole. Mais cette inversion requiert un modèle articulatoire que nous avons cherché à établir à partir de données. Nous avons mis au point un système d’acquisition de données multimodales combinant le son de parole, un flux stéréovision de la face à haute fréquence (200 images par seconde), des capteurs de localisation électromagnétiques (100Hz) collés sur la langue et la tête, ainsi qu’un flux d’images échographiques de la langue, le tout couplé à des images statiques d’IRM 3D acquises pendant des sons tenus par le locuteur. Ce système, nommé ARTIS, a été décrit dans [6]. Des méthodes ont été développées pour caractériser chaque sous-système en termes de précision spatiale, les synchroniser [29], les recalculer en 3D [28], et en traiter les données, notamment pour segmenter la langue dans les images échographiques [27] ou en IRM [23]. Ce système est non seulement le plus complet en termes de données, mais il produit également des données fondées, à savoir que les erreurs qu’elles contiennent, et la manière dont elles sont propagées d’une modalité à l’autre, sont connues [49]. Nous avons ainsi pu décrire une méthode d’estimation automatique des paramètres articulatoires chez un locuteur [25].

Ce projet est ainsi un exemple du transfert possible de nos compétences méthodologiques à des domaines applicatifs non médicaux.

Ce travail a été réalisé dans le cadre de la thèse de Michaël Aron [Aro09] (soutenue en 2009, dirigée par Marie-Odile Berger et que j’ai co-encadrée), ainsi que le séjour post-doctoral de Ting Peng (2008-2009) que j’ai co-encadrée avec Marie-Odile Berger. Cette question a été abordée dans le cadre du projet européen ASPI (2005-2008, coordonné par Yves Laprie) et du projet ANR ARTIS (2009-2012, coordonné par Yves Laprie).

1.2.4 Aspects méthodologiques

La réalisation de ces projets a requis d’aborder un ensemble de thématiques scientifiques qui leur sont transversales. Nous en proposons ici une lecture sur une grille à trois cases, reconnaissant par avance le côté parfois artificiel d’une telle classification, en particulier dans notre environnement fortement marqué par les applications.

La conception, la calibration et la qualification de systèmes d’acquisition de données.

L’objectif est ici la maîtrise des données manipulées, ainsi que leur acquisition en suivant des protocoles expérimentaux précis. Nos contributions concernent la conception de systèmes d’acquisition de données [3, 6, 17, 31], la caractérisation statistique de l’erreur attachée à un système complexe [29, 31, 49], la modélisation géométrique du capteur image [33, 41, 51] ou encore le design de protocoles d’acquisition [14, 28]. Ce souci de maîtrise et de connaissance des données se retrouve aussi dans la validation d’une imagerie pour une tâche clinique [8, 10, 35, 36, 37, 54, 55, 56]

Le recalage et la fusion de données multimodales. La multimodalité est sous-tendue par l’idée que chaque type de données informe sur une partie de l’objet d’étude. Les données médicales auxquelles nous nous intéressons sont localisées dans l’espace et le temps. Nous avons donc abordé la question du recalage spatial par le biais de capteurs [6, 28], la conception et l’usage de fantômes multimodaux [6, 11, 31], un modèle géométrique calibré par des méthodes issues de la vision robotique [30, 33, 41, 51] ou par un recalage du contenu iconique ou structurel entre images ou avec un modèle géométrique dans un contexte rigide [39, 40, 55], articulé [16] ou non-rigide dans un objectif de segmentation [23, 27, 32, 38], ou de suivi [7, 18, 19, 22]. Nous avons aussi proposé des méthodes de recalage temporel, c’est-à-dire de synchronisation, basées sur des protocoles expérimentaux [6, 29]. Enfin, la question de la visualisation de données recalées a été effleurée, à travers le concept de fluoroscopie augmentée [1, 51].

La segmentation, la reconstruction et la modélisation. Ces trois grands thèmes sont liés par leur rôle essentiel dans l’exploitation et l’analyse des données. Nous avons développé des méthodes de segmentation vasculaire [5, 21, 32, 35, 38, 50, 54], de la langue [23, 27], ou de manière plus anecdotique de scanners de troncs d’arbres [9]. Le résultat de la segmentation peut dans certains cas être assez directement utilisé, par exemple pour enrichir une visualisation. Plus souvent, une étape supplémentaire de reconstruction est requise. Nos contributions adressent la reconstruction visuelle de formes [54] ou de courbes [15], ainsi que la reconstruction par ajustement d’un modèle implicite de vaisseaux sanguins [2, 5, 20]. Outre les modèles géométriques, nous avons également développé des modèles statistiques à base d’analyse en composantes principales d’ensembles de contours à des fins de validation [8, 35, 36, 37, 54], pour guider un processus de segmentation [23] ou pour piloter un modèle articulaire [6, 25]. Un troisième type de modèle que nous avons exploré est celui des modèles physiques et plus précisément mécaniques. Le recalage et la segmentation permettent de confronter à des données réelles des instances du modèle obtenues par simulation numérique, ce qui permet d’en estimer les paramètres mécaniques [16, 24, 26]. Mes recherches actuelles s’intéressent au problème plus largement posé de guider une simulation physique interactive par des images interventionnelles.

Un thème encore plus transversal est celui de la validation, présente dans chaque publication. Celle-ci demande une collaboration proche avec les médecins, mais aussi, et en conséquence, la capacité de développer des prototypes logiciels accueillant les implantations de nos algorithmes. Nos algorithmes sont implantés dans la librairie de l’équipe Magrit, nommée RAlib [70], certains autres dans la plateforme Sofa¹, et nous avons développé une librairie logicielle nommée POLAR [13, 69] qui permet de créer rapidement des prototypes d’applications graphiques de réalité augmentée et de visualisation 3D interactive, en particulier pour les images médicales.

1. <http://sofa-platform.org>

1.3 Plan du mémoire et Guide de lecture

Ce mémoire présente de manière synthétique une sélection de travaux réalisés depuis ma thèse en 2000, et plus précisément mon arrivée à l’Inria en 2001. Le choix des thèmes abordés s’est porté sur des applications qui d’une part permettait de couvrir l’ensemble des méthodologies listées ci-dessus, et d’autre part avait fait l’objet d’un travail de thèse par au moins un doctorant que j’ai co-encadré. Pour ces raisons, et par souci de concision, les travaux en lien avec la coelochirurgie ne seront pas abordés dans ce mémoire.

Le chapitre 2 traite de nos travaux en acquisition et traitement de données articulatoires. Ces travaux ont été réalisés dans le cadre des projets européen ASPI² (12/2005-01/2009) et ANR ARTIS (01/2009-06/2013). De nombreux résultats reproduits dans ce chapitre ont été obtenus par Michaël Aron lors de sa thèse (soutenue en 2009 [Aro09]), ainsi que par Ting Peng lors de son séjour post-doctoral dans l’équipe Magrit en 2009.

Le chapitre 3 est le premier des deux chapitres qui rapportent nos travaux dans le domaine de la neuroradiologie interventionnelle. On retrouvera dans ce chapitre les résultats de notre collaboration avec GE Healthcare et le CHU de Nancy, qui ont principalement porté sur l’intégration de la 3DRA en salle d’opération par la fusion avec les images d’angiographie et de fluoroscopie. La plupart des résultats ont été obtenus lors des thèses de doctorat de René Anxionnat [Anx03] (soutenue en 2003), Sébastien Gorges [Gor07] (thèse Cifre soutenue en 2007) et Charlotte Delmas [Del17] (thèse Cifre en cours de finalisation).

Le chapitre 4 présente un deuxième pan de nos activités en lien avec la neuroradiologie interventionnelle, accomplies en collaboration avec les équipes MIMESIS et DEFROST, et le CHU de Nancy. Les résultats principaux de ce chapitre se rapportent à un modèle de surface vasculaire adapté à la simulation physique interactive, ainsi qu’au développement et à la validation de cette simulation en vue de la rapprocher quantitativement des données images interventionnelles. L’essentiel des résultats ont été obtenus dans le cadre de la thèse d’Ahmed Yureidini [Yur14] (soutenue en 2014), ainsi que du stage de Master de Pierre Glanc [Gla09] (2009). Ils ont été soutenus par les projets SIMPLE³ (*SIMulation for PLanning the Embolization of intracranial aneurisms*), Action de Recherche Coopérative (ARC) Inria que j’ai coordonnée (2006-2008) ; SOFA, Action d’Envergure Nationale (AEN) Inria, et IDeaS (*Image-Driven Simulation*), projet ANR jeune chercheur (2012-2016).

Chaque chapitre suit le même canevas en proposant une courte introduction au contexte à la fois applicatif et scientifique, puis nos principales contributions, pour se terminer sur des perspectives spécifiques aux travaux revus dans le chapitre. Le chapitre 5 conclut ce manuscrit par une évocation de nos directions de recherche à plus long terme.

Les références aux publications de l’auteur sont indiquées par un nombre. Les autres publications, de la bibliographie générale, sont référencées par les initiales des auteurs et l’année de publication.

2. *Audiovisual to articulatory SPeech Inversion*, Projet FET numéro IST-2005-021324. Voir <http://aspi.loria.fr> pour plus d’informations.

3. <http://simple.loria.fr>

Chapitre 2

Acquisition et traitement de données articulatoires

Sommaire

2.1	Introduction	15
2.1.1	Contexte scientifique	15
2.1.2	Contexte applicatif	17
2.1.3	Contributions	19
2.2	Le système ARTIS : description et principe	19
2.2.1	Présentation générale	19
2.2.2	Echographie	20
2.2.3	Capteurs électromagnétiques	20
2.2.4	Acquisition stéréovision	22
2.2.5	PC de contrôle	22
2.3	Recalage et synchronisation des données EM et US	22
2.3.1	Données et repères	23
2.3.2	Utilisation d'un fantôme	24
2.3.3	Procédure de calibration avec notre fantôme	25
2.3.4	Synchronisation EM/US/audio	28
2.4	Acquisition et recalage des données IRM	30
2.4.1	Contexte bibliographique	30
2.4.2	Protocole IRM proposé	31
2.4.3	Recalage des images IRM	34
2.5	Quelle fiabilité pour ces données ?	35
2.5.1	Évaluation visuelle	35
2.5.2	Récapitulatif des repères et transformations	36
2.5.3	Évaluation quantitative	37
2.6	Exemple d'exploitation du système ARTIS	38
2.7	Recalage et segmentation automatique d'images IRM	39
2.7.1	Hypothèses et données	39
2.7.2	Méthode	41
2.7.3	Résultats	43
2.8	Conclusion et Perspectives	43

2.1 Introduction

2.1.1 Contexte scientifique

La production de la parole est un phénomène extrêmement difficile à modéliser. Une des raisons en est que le conduit vocal est une cavité à la géométrie complexe et variable. Il est de

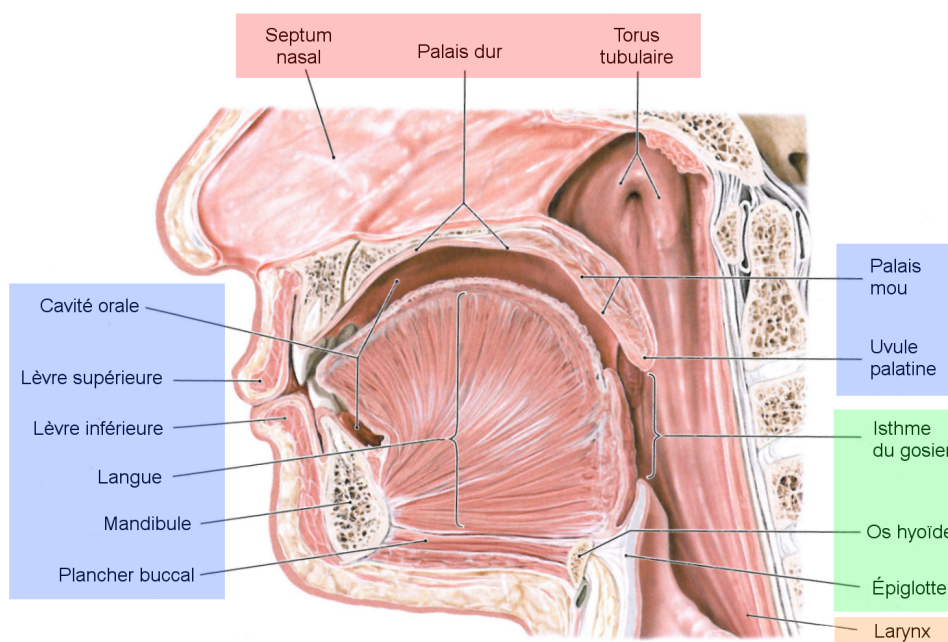


FIGURE 2.1 – Schéma du conduit vocal. En rose les fosses nasales, en bleu la cavité buccale, en vert le pharynx et en orange le larynx. D’après [SSS⁺06].

plus constitué à la fois d’organes mous (e.g. langue, luette, lèvres) et durs (e.g. dents, palais) en interaction. Une onde sonore s’y propage avec pour résultat le signal de parole. Une connaissance approfondie du conduit vocal aide à mieux aborder cette complexité géométrique, physique et mécanique. C’est pour cela que des avancées technologiques portant sur l’acquisition de données articulatoires ont souvent été la clé de progrès scientifiques en modélisation de la production de la parole. Les travaux exposés dans ce chapitre étudient la forme géométrique du conduit vocal.

Le conduit vocal peut être découpé en quatre zones d’intérêt (Figure 2.1) : la cavité buccale, les fosses nasales, le pharynx et le larynx. Il permet d’assurer les fonctions de déglutition, de respiration, et aussi celle de phonation (i.e. la production d’un son), qui nous intéresse plus particulièrement ici. Les différents éléments qui entrent en jeu dans la phonation sont appelés des articulateurs. Des détails et les références adéquates sont donnés dans le manuscrit de thèse de Michaël Aron [Aro09], et nous ne les reprendrons pas ici. Il est cependant important de savoir que trois articulateurs sont particulièrement déformables et mobiles : la langue, le palais (notamment le velum dans sa partie postérieure) et les lèvres.

Ces articulateurs sont à la fois externes et internes. Les observer est donc difficile. Par ailleurs, les méthodes d’investigation doivent être aussi peu invasives que possible car même de petits obstacles dans le conduit vocal peuvent perturber fortement la phonation. Enfin, une haute résolution est requise sur les mesures. Résolution temporelle, puis qu’on estime qu’une fréquence minimale de 60 Hz et jusqu’à 1000 Hz est nécessaire pour échantillonner sans perte le mouvement d’un articulateur (notamment pour les consonnes plosives, comme le /p/) [MBE⁺06]. Résolution spatiale également puisqu’une différence subtile de position peut faire passer d’une constriction¹ forte à une occlusion, avec un impact majeur sur le son prononcé.

Dès le début des années 1940, Chiba et Kajiyama [CK41] cherchent à exploiter de manière simultanée plusieurs sources de données, dont l’imagerie médicale, combinant photographie par rayons X, palatographie et observation laryngoscopique. Des temps d’acquisition trop longs restreignent alors l’étude aux sons émis en positions statiques (par exemple les voyelles dont on peut maintenir la phonation). Or on sait que ces positions sont non naturelles, en quelque sorte forcées et excluent du périmètre de l’étude des phénomènes dynamiques comme la coarticulation². Ces études statiques se sont poursuivies néanmoins, majoritairement à partir de données

1. Localisation du passage le plus étroit dans le conduit vocal.

2. En phonologie, un phonème est un contraste phonétique qui devient porteur de sens dans une langue. Il est

IRM, tirant ainsi profit des grands progrès réalisés vers une réduction du temps d'acquisition des IRM [BGGN91, STH96, BBR⁺02].

La cinéradiographie³ a été abandonnée suite au risque d'irradiation pour le locuteur volontaire mais également car la position des articulateurs est parfois confuse à cause des superpositions dans les images. De telles données acquises sur des corpus importants ont toutefois permis d'envisager l'hypothèse que la phonation pour un locuteur donné était commandée par un faible nombre de paramètres [Mae79], donc potentiellement étudiables à partir de données éparées. Une direction de recherche s'est donc intéressée aux capteurs de mouvements électromagnétiques (EM) [PCS⁺92], et a donné naissance à un ensemble de systèmes dédiés appelés articulographes, ou EMA (pour *ElectroMagnetic Articulograph*) : un des plus connus, dont nous disposons au Loria, est l'AG500 de chez Carstens⁴. Les capteurs actuels sont suffisamment petits (quelques millimètres) pour être collés sur les lèvres, la langue, le palais, voire la luette et peuvent donc offrir quelques points de mesure (une dizaine) sur les articulateurs avec une haute fréquence d'échantillonnage (200Hz). En revanche, le faible nombre de capteurs utilisables, associé à des difficultés pour les placer à des endroits stratégiques, ne permet de retrouver qu'une forme grossière de la langue. De plus, le signal électrique codant leur position est transmis par des fils qui perturbent la phonation [KBS06]. En parallèle, les travaux pionniers de Stone [SD95] se sont penchés sur l'utilisation de l'échographie (US pour *UltraSound*), seule modalité d'imagerie médicale capable de produire des images à une fréquence adaptée à l'étude du conduit vocal. Néanmoins, seule une partie de la langue peut être imagée à cause de nombreux obstacles (os, air), qui bloquent la transmission des ultrasons, et d'importants problèmes de répétabilité et de cohérence des données se posent, entre locuteurs ou entre deux sessions chez un même locuteur. À l'époque où les travaux exposés dans ce chapitre ont débuté, l'idée d'employer des systèmes multimodaux, combinant divers capteurs (microphone, capteurs de position EM, vidéo du visage, US, IRM, etc.), avait fait son chemin et quelques systèmes commençaient à apparaître, associés à des corpus⁵ [WH00, BEG03, Sto05, WTO⁺05, BEB⁺07]. Si le problème de la répétabilité des données, entre locuteurs différents ou sur un même locuteur entre deux sessions d'acquisition, était clairement identifié et abordé dans la littérature, les questions de fiabilité et de variabilité des données (par exemple, liées aux imprécisions intrinsèques aux images, aux artéfacts, ou encore à l'incertitude sur le détournement manuel des contours de la langue) étaient en général largement sous-estimées dans les travaux.

2.1.2 Contexte applicatif

L'acquisition de données articulatoires vise principalement à élaborer et valider des modèles du conduit vocal. En modélisation à partir de données, il est très fréquent que plusieurs capteurs, plusieurs modalités d'acquisition, soient nécessaires pour fournir un ensemble de données autorisant une caractérisation univoque du modèle. Des hypothèses, liées à la fois au modèle et à l'application, doivent alors être posées. Notre objectif ici n'est pas d'en faire une présentation exhaustive (on en trouvera un exposé assez complet dans [Aro09]), mais de permettre au lecteur de se faire une idée des types de modèles visés et des applications envisagées pour mieux juger de la pertinence des choix faits pour concevoir le système ARTIS présenté dans ce chapitre.

Quelques applications des modèles articulatoires

Les modèles articulatoires sont liés à la production de la parole. Dans le sens direct, la principale application en est la synthèse acoustique. On sait générer un signal acoustique à partir

susceptible d'être prononcé de façon différente selon les locuteurs mais aussi selon sa position et son environnement au sein du mot : c'est le phénomène de coarticulation.

3. Imagerie par rayons X à fréquence d'acquisition de jusqu'à 30 images par seconde.

4. <http://www.articulograph.de>

5. Ensemble de phonèmes et de phrases, fixé par la communauté, entraînant des mouvements de locution représentatifs de l'espace articulatoire. Un corpus permettant d'explorer l'intégralité de l'espace articulatoire associé à une langue est en général long à acquérir et de nombreuses bases de données se restreignent à une partie de cet espace.

de la fonction d'aire [Fan60], qui décrit le conduit vocal comme un tube droit mais à section variable. Une manière classique de déterminer cette fonction calcule la largeur du conduit vocal en tout point de sa ligne centrale, à partir de ses contours détournés dans le plan médio-sagittal.

À l'inverse, l'inversion acoustique articulatoire vise à récupérer automatiquement l'évolution temporelle de la forme du conduit vocal à partir du signal acoustique. Ce problème est extrêmement difficile, d'autant plus qu'il est sans nul doute mal posé. En effet, il existe un consensus pour reconnaître que la compréhension de la parole en exploite la multimodalité : la vision des articulateurs (position de la mâchoire, des lèvres, visibilité des dents ou de la langue) améliore l'intelligibilité de la parole. Nous en avons tous fait l'expérience en milieu bruyant, et les malentendants s'en servent pour lire sur les lèvres. Un modèle est nécessaire ici car le problème de l'inversion acoustique articulatoire est connu pour n'être pas inversible. Il n'existe pas de carte univoque entre les domaines acoustique et articulatoire : un grand nombre de formes du conduit vocal peuvent produire le même son. Un des objectifs de l'acquisition des données est d'améliorer notre connaissance des formes admissibles pour le conduit vocal afin d'en tirer des contraintes à la fois réalistes et suffisamment restrictives.

La résolution de l'inversion acoustique articulatoire ouvrirait la voie à des applications dans des domaines divers. Engwall [Eng08] a montré qu'un retour visuel réaliste sur le positionnement de la langue permet à un individu français d'améliorer sa prononciation du suédois. On peut alors songer à des têtes parlantes virtuelles reproduisant les positions et mouvements des articulateurs du conduit vocal. Outre l'apprentissage des langues, cela pourrait bénéficier à l'orthophonie en guidant les personnes malentendantes dans leur articulation. D'autres utilisations des têtes parlantes sont aussi envisageables dans le domaine du cinéma d'animation. Dans le domaine médical, elles pourraient aider les patients ayant subi une glossectomie (ablation partielle de la langue) à réapprendre à parler [BTB⁺05]. De manière pro-active, le modèle articulatoire pourrait aider à prédire les difficultés de parole et guider ainsi certains choix opératoires ou encore être à la base de la conception de prothèses adaptées aux mouvements de langue possibles, ainsi qu'au patient.

Quelques modèles articulatoires

Un modèle articulatoire se présente en général géométriquement comme une coupe médio-sagittale du conduit vocal, à partir de laquelle il est possible de tirer une modélisation acoustique. Outre le modèle à fonction d'aire que nous venons d'évoquer, deux grandes familles de modèles ont été développées.

Les modèles biomécaniques ont pour ambition d'intégrer un maximum de données physiologiques et géométriques sur les articulateurs, et de modéliser leurs contacts entre eux (par exemple entre la langue et le palais) et avec leur environnement (os, muscles...). Les premiers modèles masse-ressort en 2D ont évolué vers la 3D et ils sont aujourd'hui exprimés sous forme d'éléments finis en 3D. Ils reposent cependant sur un très grand nombre de paramètres les rendant complexes, coûteux en temps de calcul et il est difficile d'assurer à la fois leur stabilité numérique et leur accord avec des données réelles. La focalisation porte aujourd'hui sur la modélisation de la langue, l'articulateur le plus mobile et complexe [BRPP15].

Les modèles statistiques sont basés sur une analyse factorielle de la forme du conduit vocal observé en locution. Peu de paramètres sont en général suffisants pour que ces modèles linéaires recouvrent en quasi-totalité l'espace des formes du conduit vocal. Par exemple, le modèle de Maeda [Mae79] permet de reproduire 96% de la variance observée sur les données avec juste 7 modes linéaires de déformation (voir aussi la thèse de Michaël Aron [Aro09] pour un exposé détaillé ou [25] pour une description rapide). Ce modèle a servi de guide aux études que nous avons réalisées dans le cadre de l'étude de la parole.

2.1.3 Contributions

Nos travaux liés à l'acquisition et aux traitements de données articulatoires se sont déroulés dans le cadre du projet européen ASPI⁶ (12/2005-01/2009) et projet ANR ARTIS (01/2009-06/2013), dont un des principaux résultats fut le système d'acquisition de données articulatoires audio-visuelles ARTIS [6]. Ce système permet d'acquérir conjointement, chez un locuteur, le son prononcé, la position de capteurs EM, une séquence US dans le plan médio-sagittal de la langue, et une séquence stéréovision du visage permettant d'en reconstruire la surface. Par ailleurs, ces acquisitions dynamiques sont complétées par des acquisitions statiques d'IRM 3D du conduit vocal. Nos contributions ont concerné l'acquisition et le traitement des données issues des modalités d'imagerie médicale ou couramment utilisées dans le contexte médical, autrement dit, les données multimodales EM, US et IRM, et la dernière modalité, acquisition par stéréovision de la face, ne sera mentionnée dans ce chapitre que de manière anecdotique.

La Section 2.2 décrit le système ARTIS et en expose le principe d'une acquisition.

La Section 2.3 reprend nos contributions principales concernant le sous-système couplant capteurs EM et images US. Nous avons caractérisé expérimentalement la précision géométrique de ce sous-système, et proposé une première méthode de calibration du recalage géométrique entre ces deux modalités [31]. Nous avons proposé une procédure de synchronisation de ce même sous-système avec le signal audio [29]. Nous décrivons dans ce mémoire les protocoles finalisés de sa calibration, en particulier une méthode de calibration géométrique du couple EM/US centrée sur un nouveau fantôme [6].

La Section 2.4 se focalise sur nos travaux autour de l'IRM. Différents protocoles d'acquisition IRM du conduit vocal en position articulatoire maintenue étaient proposés dans la littérature. Nous les avons comparés expérimentalement grâce au système ARTIS. Nous en avons déduit un protocole d'acquisition IRM original [28] que nous avons suivi pour acquérir un premier corpus de plus de 20 minutes. Une première expérience d'inversion articulatoire audio-visuelle a pu être menée [25]. Nous avons également mis au point une méthode de segmentation du contour de la langue dans les images IRM du plan médio-sagittal de différents locuteurs, basée sur un modèle statistique de forme appris chez un locuteur de référence [23].

Les principales publications sur lesquelles est basé ce chapitre sont : [6, 23, 25, 27, 28, 29, 31]

De nombreux résultats reproduits dans ce chapitre ont été obtenus par Michaël Aron lors de sa thèse (soutenue en 2009 [Aro09]). L'algorithme de segmentation de la langue en IRM a été le produit du séjour post-doctoral de Ting Peng en 2009. J'ai co-encadré Michaël et Ting avec Marie-Odile Berger.

2.2 Le système ARTIS : description et principe

Nos travaux ont eu pour objectif de mettre au point un système d'acquisition intégrant plusieurs sources de données complémentaires pour assurer une meilleure couverture des articulateurs. Par ailleurs, un objectif général visait à proposer un système peu coûteux (comparé aux articulographes par exemple) et surtout aisément reproductible par d'autres groupes de recherche. Enfin, un souci constant a été de fiabiliser ces données par une étude systématique des sources d'erreur et la mise en œuvre de traitements automatiques.

2.2.1 Présentation générale

Le système ARTIS permet l'acquisition simultanée de données audio (sons prononcés), EM (capteurs de position sur les articulateurs), US (contour supérieur du dos de la langue) et stéréovision (surface du visage). Ces données sont synchronisées et exprimées dans le même repère spatial. Par ailleurs, il intègre une procédure d'ajout de données IRM dans ce même repère.

La Figure 2.2 montre une photo du système d'acquisition ARTIS dans sa globalité, accolée à une prise de vue pendant son utilisation. On remarquera en particulier les croix vertes superposées

6. *Audiovisual to articulatory Speech Inversion*, Projet FET numéro IST-2005-021324. Voir <http://aspi.loria.fr> pour plus d'informations.

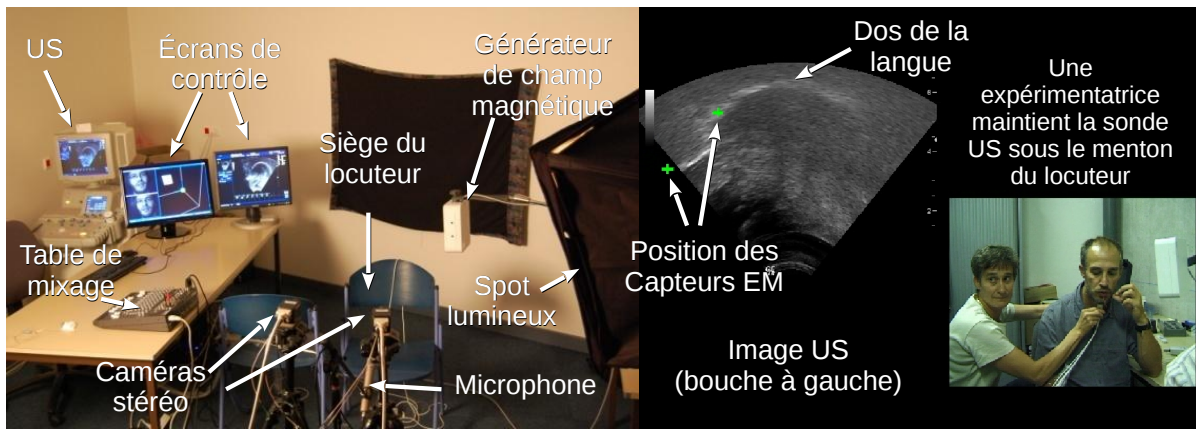


FIGURE 2.2 – Le système ARTIS : (à gauche) photo du système dans sa globalité présentant les divers éléments le composant ; (au centre) exemple d'image US sur laquelle ont été superposées des croix vertes indiquant la position des capteurs EM (à gauche sur l'apex, à droite sur le dos de la langue) (à droite) pendant l'acquisition, le locuteur est placé dans le champ du générateur EM et une expérimentatrice maintient la sonde US sous le menton du locuteur, tout en vérifiant la qualité de son placement sur l'écran de contrôle afin d'assurer une image dans le plan médio-sagittal.

à l'image échographique au centre, qui indiquent les positions des capteurs EM collés sur la langue. La Figure 2.3 reproduit un schéma des différents composants formant l'architecture du système ARTIS.

Les différents éléments le composant ont été choisis selon deux critères principaux : 1) données d'une qualité adaptée à la capture de mouvements articulatoires ; et 2) disponibilité dans une offre commerciale standard afin que le système puisse être reproduit par d'autres équipes de recherche.

2.2.2 Echographie

Le premier élément du système est un échographe Logiq5 (GE Healthcare, the Chalfont St Giles, UK). Le choix d'un échographe fixe, et donc encombrant, a été fait car la qualité image était nettement supérieure à celle proposée par les échographes portables. Par ailleurs, le Logiq5 permettait l'enregistrement et le transfert des images US au format numérique Dicom, donc sans perte contrairement aux captures analogiques faites par d'autres systèmes d'acquisition [Sto05].

La sonde microconvex 8C utilisée produit des ultrasons dans la plage de fréquence 5-9 MHz. Sa forme incurvée permet de garder un bon contact avec la peau sans gêner la locution. Le protocole d'acquisition mis au point offre une bonne couverture de la langue avec une zone image d'une largeur de 8.6 cm et une profondeur de 7 cm. Les images, de taille 532×434 pixels, sont acquises à une fréquence de 66 images par seconde pour une taille pixel annoncée de 0.1753 mm/pixel.

Le signal ultrasonore est aisément bloqué par l'air ou l'os hyoïde. Il est donc impossible de visualiser le palais ou encore l'extrémité de la langue (l'*apex*) en échographie. Une autre difficulté liée à ce système en particulier vient du fait que l'acquisition se commande par un signal qui en indique la fin : les 15 secondes précédant ce signal sont alors enregistrées.

2.2.3 Capteurs électromagnétiques

Le second élément du système ARTIS est un système électromagnétique de capture du mouvement. Les articulographes sont des systèmes très coûteux et, au moment où ce travail a débuté, ils ne permettaient pas d'accéder aux données en temps réel et gênaient l'acquisition de la face par stéréovision. Si ces deux derniers inconvénients ont été résolus depuis, le prix reste toujours très élevé. Nous avons donc opté pour le système proposé par la société Northern Digital Inc.

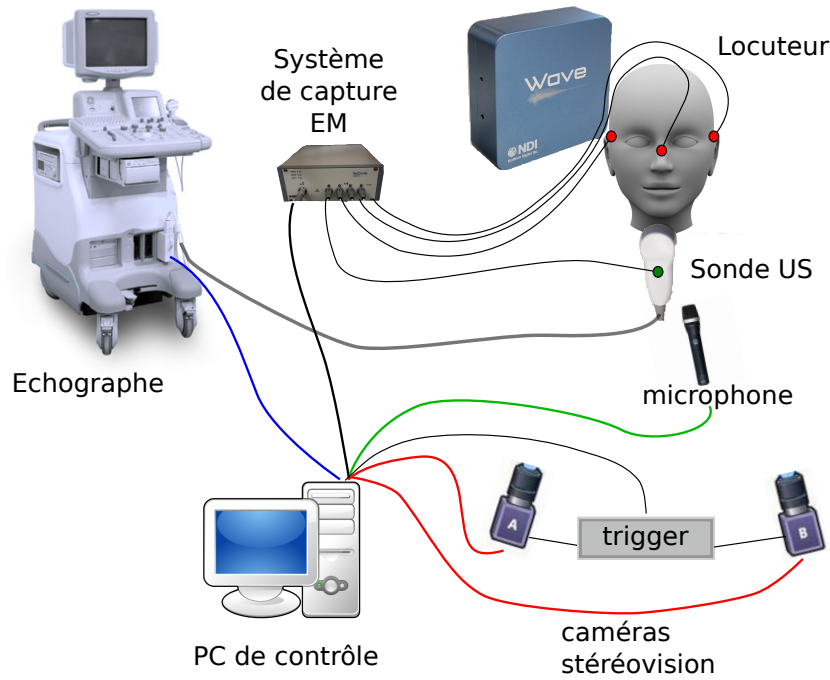


FIGURE 2.3 – Architecture du système ARTIS.

(NDI, Waterloo, Ontario, Canada). Nos premières expériences ont été réalisées avec le système Aurora, mais NDI produit depuis 2009 le système Wave⁷, dédié à l'acquisition du mouvement articulaire, et dont le développement, sans aller jusqu'à prétendre à la moindre once de paternité, a sans nul doute été nourri par les échanges réguliers que nous avons eus, en particulier Michaël Aron, avec les ingénieurs de chez NDI. La version actuelle du système ARTIS intègre le système Wave, auquel ce texte va donc uniquement s'attacher.

Le système Wave permet de capturer le mouvement de 16 capteurs à 5 degrés de liberté (*DDL*, 3 en position et 2 en orientation, pas de mesure de rotation autour de l'axe principal du capteur), ou 8 capteurs à 6 DDL (2 capteurs 5 DDL font un capteur 6 DDL) à une fréquence de 100 Hz (200 et 400 Hz sont disponibles à condition de réduire le nombre de capteurs, ce que nous n'avons pas choisi de faire). Un capteur a une taille de 2×3 mm. La position et l'orientation d'un capteur sont déduites des caractéristiques du courant induit par sa plongée dans le champ EM produit par un générateur de champ magnétique (voir Figure 2.2). Cette mesure n'est valable que dans un volume cubique de côté 50 cm situé en avant du générateur.

Trois capteurs 5DDL sont placés sur la tête du locuteur (deux derrière les oreilles et un en haut de l'arête du nez, voir Figure 2.3) : ils définissent un repère spatial, fixe par rapport à la tête du locuteur. Ce repère, que nous noterons \mathcal{R}_0 , définit le repère de référence pour le système ARTIS. Deux capteurs 5DDL sont collés sur la langue, dans le plan médio-sagittal : un sur le dos de la langue et l'autre sur son apex, point important dans la locution, mais invisible sur les images US. Par ailleurs, un capteur 6DDL est collé sur la sonde US afin d'en suivre le mouvement. Enfin, un capteur 6DDL est disponible pour localiser l'extrémité d'un stylet électromagnétique. Ce dernier capteur est utilisé dans le recalage de la surface stéréo du visage ou de l'IRM dans le repère EM, ainsi que pour le calibrage du recalage EM/US décrit plus en détail dans la Section 2.3.

Nous avons vérifié expérimentalement que la mesure indiquée par les capteurs EM comprend une erreur en translation inférieure au millimètre et en rotation inférieure à 0.5° à condition que les capteurs ne s'approchent pas des limites de la zone de validité. La sonde US ne perturbe pas les mesures de manière significative [6, 31].

7. Voir <http://www.ndigital.com/msci/products/wave-speech-research/>

2.2.4 Acquisition stéréovision

Le troisième élément est formé par une paire de caméras JAI-Pulnix TM-6740CL (JAI, Valby, Denmark). Ces caméras sont configurées pour acquérir des séquences d'images en niveaux de gris, de taille 640×480 pixels à 198 images par seconde. Elles sont synchronisées grâce à un trigger externe (CC320 Machine Vision Trigger Timing Controller, Gardasoft Vision, Cambridge, UK). À cette fréquence, le temps de pose pour une image est plus court que les variations d'intensité lumineuse que connaissent les néons qui éclairent la salle d'expérimentation. Nous avons donc employé deux spots lumineux Super Cool-Lite 9 (interFit Lighting) comme seul éclairage.

Je me suis intéressé à la reconstruction du visage par stéréovision en tant que partie du stage de DEA de Sébastien Demanges que j'ai encadré en 2004 [Dem04]. La méthode développée estime une carte de disparité par maximisation du score de corrélation locale. La méthode applique une résolution multi-échelle au cours de laquelle ce score est identifié au poids d'une coupe dans un graphe valué (*graph-cuts*) [KZ01, KZ04]. Bien que précise, cette méthode ne permet cependant qu'une reconstruction de la surface du visage, sans cohérence temporelle des sommets du maillage. Ce problème classique d'ouverture empêche de suivre un point physique sur cette surface lors de sa déformation, et donc de mesurer correctement le mouvement d'un articulateur. Une technique alternative de suivi, reconstruction puis maillage d'un ensemble de marqueurs peints sur le visage a donc été mise en œuvre [WDBP⁺05] sans que je participe à son développement.

Des marqueurs sont peints sur les lèvres et la mâchoire inférieure pour capturer le mouvement de ces articulateurs. D'autres sont peints sur le front et l'arête du nez pour définir un repère fixe par rapport à la tête du locuteur. Enfin, des marqueurs supplémentaires sont placés afin d'échantillonner de manière suffisamment fine la surface du visage (quelques millimètres à centimètres d'écart).

2.2.5 PC de contrôle

Les différentes modalités, y compris l'enregistrement du signal audio sur une table de mixage, sont centralisées par un PC de contrôle qui utilise 4 disques RAID0 afin d'atteindre une vitesse d'enregistrement de 360 Mo/s, nécessaire au stockage en ligne des séquences stéréo [6].

Un logiciel graphique a été développé en Python pour vérifier la bonne position de la sonde US lors de l'enregistrement des données. Son principe tire profit, d'une part de la disponibilité en temps réel des données EM, qui permet d'afficher la position des deux capteurs 5DDL placés sur la langue, et d'autre part du calibrage EM/US (voir la Section 2.3 ci-dessous), associé au suivi de la sonde US par le capteur 6DDL qui y est collé. Le plan image US peut ainsi être représenté dans le même repère que les deux capteurs de la langue et il est aisé de vérifier qu'il en est proche, faute de quoi une alarme est émise et la correction manuelle est intuitive.

Ce logiciel est affiché sur un écran de contrôle que surveille en parallèle de l'écran de l'échographe le manipulateur chargé de tenir la sonde US. Un autre écran permet au locuteur de lire les sons et phrases du corpus à prononcer.

2.3 Recalage et synchronisation des données EM et US

Comme dans tout système multimodal, un préalable à la fusion et l'exploitation conjointe de données EM et US est leur expression dans un même repère spatial, autrement dit leur recalage. De plus, les données sont ici dynamiques, ce qui nécessite en plus de les exprimer dans un même repère temporel, c'est-à-dire les synchroniser. Cette section présente les méthodes que nous avons développées pour résoudre ces deux problèmes, en débutant par le recalage spatial.

Note préliminaire : nous traitons ici de recalage rigide, qui s'exprime comme un changement entre deux repères orthonormés \mathcal{R}_A vers \mathcal{R}_B . Le recalage revient donc à trouver une transformation rigide, que nous noterons ${}^B T_A$ et qui sera exprimée en coordonnées homogènes dans l'espace projectif à trois dimensions par la matrice 4×4 . Une telle transformation est paramétrée par 6

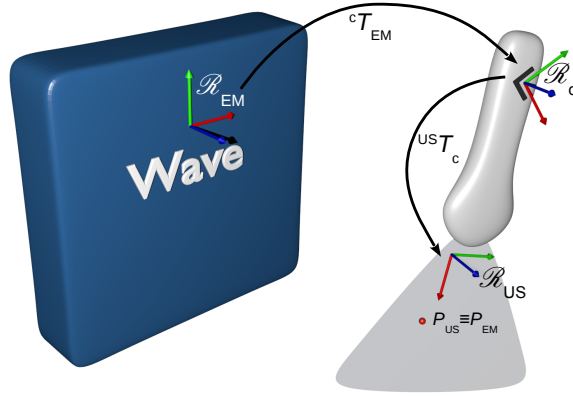


FIGURE 2.4 – Repères spatiaux et transformations utilisés pour le recalage EM/US. Le capteur 6DDL est indiqué par la petite équerre noire sur la sonde US. Un point visible dans l'image US, et donc exprimé dans \mathcal{R}_{US} , correspond physiquement à un point 3D dans \mathcal{R}_{EM} . La matrice S_{US} permettant de passer des coordonnées métriques à des coordonnées pixel n'est pas représentée sur la figure (voir Eq. 2.2).

paramètres (3 angles de rotation et 3 translations).

$${}^B T_A = \begin{pmatrix} \mathbf{R} & t \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.1)$$

où \mathbf{R} est une matrice de rotation, et t un 3-vecteur. Une transformation rigide est toujours inversible. Nous utiliserons le point («.») pour désigner l'action d'une transformation, sur un point, une droite, un plan ou sur une autre transformation (composition).

2.3.1 Données et repères

Les données EM se présentent sous la forme d'un 3-vecteur et d'un quaternion, le premier donnant la position du capteur et le deuxième donnant son orientation par rapport au générateur de champ magnétique. Mathématiquement, cela implique l'existence d'un repère lié au générateur de champ magnétique (noté \mathcal{R}_{EM}) et, pour chaque capteur, d'un repère local (noté \mathcal{R}_c , c pour capteur). Il est important de noter que le générateur de champ magnétique est positionné manuellement avant chaque acquisition : le repère \mathcal{R}_{EM} n'est donc pas immuable d'une acquisition à l'autre. Chaque donnée EM se traduit alors comme la transformation rigide ${}^{EM}T_c$. Cette transformation est complètement spécifiée dans le cas d'un capteur 6DDL, par exemple celui qui est collé sur la sonde US qui nous intéresse particulièrement ici. Les positions dans \mathcal{R}_{EM} et \mathcal{R}_c sont indiquées en millimètres.

Les données US sont des images, ou plus précisément chaque pixel d'une image. Suivant [PRGB98], on définit de manière arbitraire un repère \mathcal{R}_{US} en indiquant que la position d'un pixel sur la colonne u et la ligne v dans ce repère est le vecteur homogène $P_{US} = (u, v, 0, 1)$.

Ce pixel quelconque a pour coordonnées le vecteur P_{EM} dans \mathcal{R}_{EM} et la relation fondamentale suivante en découle (voir la Figure 2.4) :

$$P_{US} = S_{US}^{-1} \cdot {}^{US}T_c \cdot {}^cT_{EM} \cdot P_{EM} \quad (2.2)$$

où S_{US} permet de passer des coordonnées en pixel aux coordonnées en millimètres. Nous faisons l'hypothèse que l'image US ne comporte aucune distorsion (ce qui n'est strictement vrai que

dans le cas où les ultrasons traversent un milieu où la vitesse du son est homogène), ce qui nous permet d'écrire $S_{US} = \text{diag}(r_h, r_v, 1, 1)$, r_h et r_v étant respectivement les tailles pixel horizontale et verticale de l'image US.

Résoudre le problème de recalage revient à déterminer les matrices S_{US} et ${}^{US}T_c$. Il est important de noter que, le capteur 6DDL étant solidaire de la sonde US, ${}^{US}T_c$ ne dépend pas de la position du générateur de champ magnétique et reste constant tant que le lien entre le capteur et la sonde est maintenu.

2.3.2 Utilisation d'un fantôme

Dans la littérature, les investigations de ce problème se sont majoritairement placées dans le cadre applicatif de l'échographie 3D main-libre. Une image 3DUS est reconstruite en agglomérant les différentes images US dans un même repère. Ce repère est en général celui dans lequel est localisé un capteur de mouvement (EM, optique ou mécanique) solidaire de la sonde US.

Les nombreuses méthodes proposées calibrent ces transformations au moyen d'un fantôme [MLLC05] ou d'un objet disponible pendant l'acquisition dont la géométrie est connue [VPOS16] et dont les structures (points, lignes, plans,...) sont visibles et facilement détectables dans une image US. Parmi les fantômes recensés par [MLLC05], on ne retrouve aujourd'hui dans la littérature que ceux basés sur deux grands principes : acquérir des images US en coupe d'une ligne [KS05, VPOS16] (ou un ensemble de lignes [CTEA09]) ou d'un plan [NAAR15] (ou un ensemble de plans), ces deux structures se retrouvant parfois dans un même fantôme [NAAR14].

Choix du type de fantôme

Chaque design a ses avantages et inconvénients. Un compromis existe néanmoins pour rejeter les fantômes dont la structure à imager est un point (ou ensemble de points) 3D à cause de la difficulté expérimentale à imager sous des incidences diverses un point donné.

Après avoir testé dans [31] un ajustement de la méthode de Khamene et Sauer [KS05], nous avons choisi de concevoir notre propre fantôme afin d'améliorer la précision de la calibration. Nous avons opté dans [6] pour un design basé sur des lignes pour deux raisons : d'une part, d'un point de vue pratique, obtenir une image US d'une ligne est facile, et d'autre part le résultat est un point qui peut se détecter de manière automatique dans l'image, et être identifié de manière univoque si les points forment un *pattern* par exemple. Par ailleurs, d'un point de vue mathématique, lignes et plans correspondent à un problème équivalent, puisqu'une droite 3D peut être vue comme l'intersection de deux plans [RTMT10, VPOS16]. Il n'y a donc pas d'intérêt théorique à préférer l'un à l'autre.

Design et caractéristiques de notre fantôme

Le design du fantôme a été guidé par quatre problèmes pratiques. En premier lieu, le son ne se propage pas dans l'air. Il faut donc que les structures à imager (droites) soit plongées dans un substrat, un fond échogène. La grande majorité des fantômes proposés emploient l'eau à 37°C dans laquelle le son se propage à 1570 m/s, une vitesse proche de celle des tissus humains (1540 m/s en moyenne). Ceci pose deux soucis : d'une part obtenir de l'eau à 37°C prend un peu de temps, du matériel un peu spécifique et du soin ; par ailleurs, il est encore plus difficile de maintenir l'eau à cette température, sauf à presser l'expérimentation, ce qui n'est jamais une bonne chose. Le design de notre fantôme est basé sur le modèle de fantôme 555 produit par la société CIRS Inc (Norfolk, VA), composé d'un ensemble de fils en nylon parallèles immergés dans un gel au sein duquel la vitesse de propagation du son est la même que dans la moyenne des tissus humains (1540 m/s).

Ensuite, les fantômes à base de lignes ont une structure ouverte (pour les remplir d'eau et y plonger la sonde US). Cela pose des problèmes de pérennité de la géométrie et une fragilité mécanique à moyen terme. Notre fantôme est un cube fermé, avec trois côtés solides et un côté plastifié sur lequel la sonde US peut être appliquée pour prendre des images.

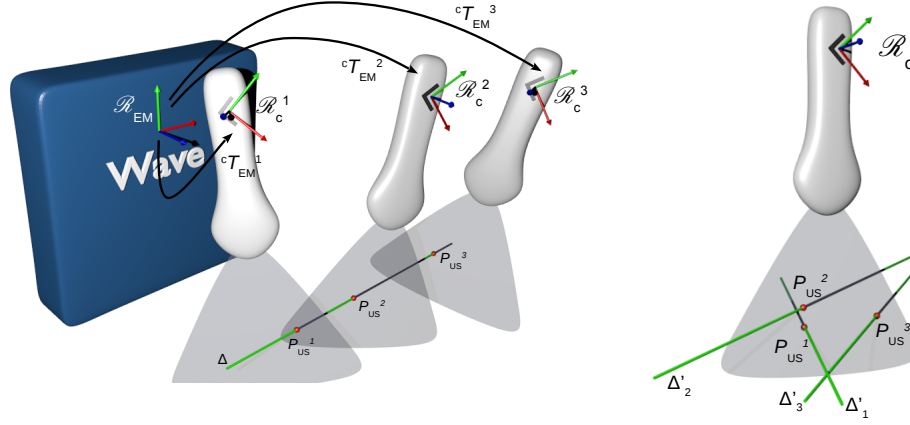


FIGURE 2.5 – Calibration à l’aide d’un fantôme à une droite Δ . À gauche : on se place dans le repère fixe \mathcal{R}_{EM} et plusieurs images de la droite sont prises. Pour chaque image $\#k$, un point P_{US}^k est détecté comme l’intersection de Δ avec le plan image, et une transformation ${}^cT_{EM}^k$ est lue sur le capteur. À droite : si on se place dans le repère mobile \mathcal{R}_c , le problème de recalage s’exprime comme la recherche de la similitude qui fait correspondre le nuage de points $\{P_{US}^k\}$ aux droites $\{\Delta_k\}$, transformées de Δ par chaque ${}^cT_{EM}^k$.

D’un point de vue mathématique, Khamene et Sauer évoquent un risque de dégénérescence de la solution si on utilise une seule ligne [KS05]. Ce problème se conçoit aisément si on reformule le problème en prenant comme repère de base \mathcal{R}_c , le repère local lié au capteur. Prendre N images US d’une seule droite Δ pour diverses positions et orientations de la sonde revient à prendre une image US de plusieurs droites $\{\Delta'_k\}_{k \in [1, N]}$, transformées de Δ par la transformation vers chaque repère local (voir la Figure 2.5). Le cas dégénéré apparaît alors comme celui pour lequel les droites $\{\Delta'_k\}$ sont parallèles (indétermination de la composante de translation parallèle à ces droites). Ce problème est intéressant car si en théorie, il est aisé d’éviter le cas dégénéré, ce n’est pas le cas en pratique où l’expérimentateur, pour avoir de belles images, a inconsciemment tendance à reproduire la même orientation du plan US, perpendiculaire à la droite. Nous avons donc opté pour un faisceau de droites. Cela permet dans un premier temps d’avoir plus de données (points P_{US}) en une seule image. Ces droites sont parallèles pour des raisons pratiques de construction. Nous avons donc ajouté un deuxième faisceau de droites, orthogonal au premier (voir la Figure 2.6). Cela n’apporte rien si les deux faisceaux sont imagés séparément (la sonde sera elle aussi naturellement tournée de 90° et cela n’apportera aucune direction nouvelle en vertu du principe de la Figure 2.5). En revanche, l’expérimentateur peut très facilement appliquer la consigne d’acquérir des images où des droites des deux faisceaux sont visibles, ce qui force à prendre des directions de droites à 45° par rapport au plan US. Chaque faisceau a été conçu pour former un pattern de points distinct, afin d’identifier le faisceau, et asymétrique, afin de ne pas introduire d’ambiguïté sur la rotation autour de l’axe vertical Y .

Enfin, il est nécessaire de pouvoir exprimer chaque droite du fantôme dans le repère \mathcal{R}_{EM} . Pour cela, nous avons demandé au fabricant de marquer par une croix blanche les deux extrémités de chaque fil sur la paroi extérieure. Ces extrémités sont localisées dans \mathcal{R}_{EM} grâce au stylet EM en début de calibration.

La Figure 2.6 donne toutes les caractéristiques de notre fantôme. Il comporte 77 fils (154 extrémités). Il a été construit selon nos indications et design spécialement pour nous par la société CIRS Inc. (Norfolk, VA).

2.3.3 Procédure de calibration avec notre fantôme

Une première étape calibre la transformation ${}^{EM}T_f$ afin de pouvoir exprimer les équations des droites dans le repère \mathcal{R}_{EM} . Dans un deuxième temps, la transformation $A = S_{US}^{-1} \cdot {}^{US}T_c$

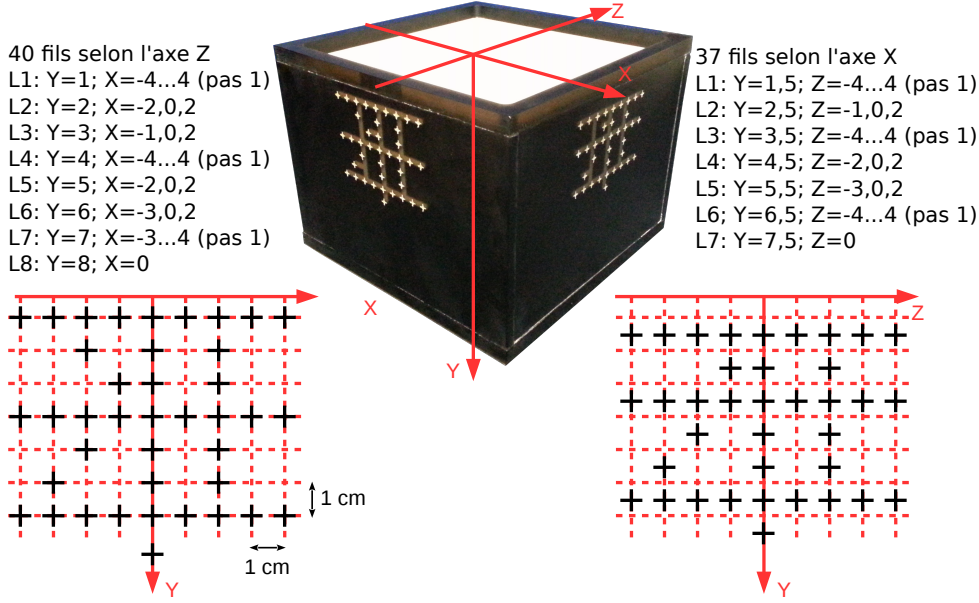


FIGURE 2.6 – Spécifications de notre fantôme pour la calibration spatiale EM/US. À gauche, les croix noires donnent les coordonnées selon X et Y des extrémités des fils parallèles à l'axe Z. Les positions sont données en centimètres, avec un espacement caractéristique de 1 cm pour la grille en pointillés rouges. À droite, les croix noires donnent les coordonnées Z et Y des extrémités des fils parallèles à l'axe X. Ils sont intercalés avec le faisceau précédent grâce à un décalage vertical de 0.5 cm. Au centre, la photo montre notre fantôme, avec une surface supérieure où placer la sonde US et les extrémités des fils marquées par des croix blanches sur ses face latérales (idem pour les deux faces cachées). Les axes X,Y et Z sont superposés en rouge à cette photo.

est calibrée pour déterminer la position du capteur EM sur la sonde US, ainsi qu'affiner la taille pixel des images US.

Calibrage de ${}^{\text{EM}}T_f$

Le système EM est installé et le fantôme est placé dans le volume de validité. Chaque extrémité de fil est ensuite pointée pendant environ un seconde avec le stylet EM afin d'acquérir un nuage de points 3D proches pour chaque extrémité (100 points par seconde). Les 154 plus gros groupes de points 3D sont extraits de tous les points acquis par classification aux plus proches voisins (distance de 1 mm, avec la contrainte d'une distance minimale de 5 mm entre groupes). Le point médian est conservé pour chaque groupe, ce qui fournit les 154 extrémités recherchées. Quatre plans sont extraits de ces points par RANSAC, puis les faces du fantôme sont identifiées par reconnaissance de patron (*pattern matching*, 4 patrons possibles), ce qui permet d'associer chaque extrémité à un numéro de ligne unique dans le modèle du fantôme. La transformation ${}^{\text{EM}}T_f$ entre le repère du fantôme \mathcal{R}_f et \mathcal{R}_{EM} est obtenue en minimisant le critère :

$$\mathcal{C}_f = \frac{1}{N_f} \sum_i d(P_{i,0}, {}^{\text{EM}}T_f \cdot \Delta_i^f)^2 + d(P_{i,1}, {}^{\text{EM}}T_f \cdot \Delta_i^f)^2 \quad (2.3)$$

où $N_f = 77$ est le nombre de fils, et $d(\cdot)$ est la distance euclidienne entre un point et une droite. $\{\Delta_i^f\}$ sont les droites exprimées dans \mathcal{R}_f (connues par design). $P_{i,0}$ et $P_{i,1}$ sont les extrémités de la droite Δ_i^f mesurées dans \mathcal{R}_{EM} . Ce critère est minimisé par l'algorithme de Powell pour déterminer les 6 paramètres de la transformation rigide ${}^{\text{EM}}T_f$. Le résidu obtenu de $\sqrt{\mathcal{C}_f} = 0.65\text{mm}$ est compatible avec la précision des capteurs EM. En conséquence, chaque droite Δ_i du fantôme peut être exprimée dans \mathcal{R}_{EM} comme $\Delta_i = {}^{\text{EM}}T_f \cdot \Delta_i^f$.

Calibrage de $A = S_{\text{US}}^{-1} \cdot {}^{\text{US}}T_c$

Une étape d'optimisation suit une initialisation manuelle.

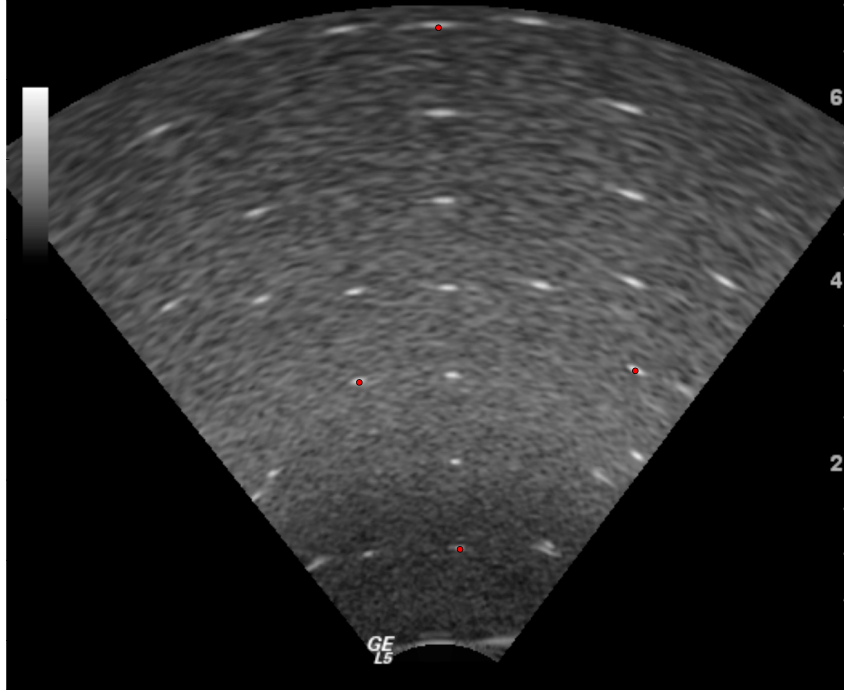


FIGURE 2.7 – Initialisation manuelle de la similitude A au moyen de 4 points (points rouges).

Initialisation : Une image US est prise de manière à ne montrer que le faisceau de droites parallèles à l'axe Z de \mathcal{R}_f . Il existe une transformation affine entre le patron montré sur la Figure 2.6 (gauche) et le dessin formé par les points dans l'image (intersection d'un plan avec un faisceau de droites parallèles). Nous identifions manuellement 4 fils dans l'image (voir la Figure 2.7) pour estimer cette transformation affine, d'où nous extrayons une estimation initiale r_h^0 et r_v^0 des tailles pixel, ainsi que la rotation et la translation dans le plan de l'image US. En supposant de plus que le plan de l'image US est perpendiculaire à l'axe Z dans \mathcal{R}_f , i.e. qu'il n'y a pas de composante rotationnelle hors du plan US, nous obtenons, à une translation selon Z près, une valeur initiale S^0 pour la similitude qui permet de passer de \mathcal{R}_f à \mathcal{R}_{US} . Nous pouvons alors initialiser A par :

$$A^0 = S^0 \cdot {}^{EM}T_f^{-1} \cdot {}^cT_{EM}^{-1} \quad (2.4)$$

où ${}^cT_{EM}$ est lue sur le capteur collé à la sonde au moment où l'image initiale est prise.

Optimisation : Plusieurs images US sont acquises dans diverses orientations, y compris des images qui montrent simultanément les deux faisceaux. Nous appliquons un fort filtre bilatéral (taille de fenêtre de 10 pixels, écart-type spatial de 10 pixels), suivi d'un seuil. Les points d'intersection entre les droites et le plan image US sont les centres de gravité des composantes connexes. La mise en correspondance points-droites se fait au plus proche voisin. Bien que le critère usuel à minimiser soit basé sur une distance euclidienne point-droite en 3D, nous avons trouvé expérimentalement une meilleure stabilité et qualité du recalage avec une distance point-point exprimée dans le plan US. Le critère que nous minimisons est donc :

$$\mathcal{C}_A = \frac{1}{N_c} \sum_k \sum_{i \in \mathcal{L}_k^*} |P_{US}^{i,k} - \Pi \cdot A \cdot {}^cT_{EM}^k \cdot \Delta_i|^2 \quad (2.5)$$

où k est l'indice de l'image, i est l'indice de la droite considérée, \mathcal{L}_k^* rassemble les indices des droites présentant une intersection avec le plan US $\#k$ étant donné l'estimation courante de A , ${}^cT_{EM}^k$ est la transformation donnée par le capteur pour l'image $\#k$ et Π est un opérateur retournant le point intersection entre une droite et le plan $z = 0$. La qualité de notre initialisation permet de minimiser ce critère avec l'algorithme de Powell afin d'estimer les 8 paramètres de A avec un résidu de $\sqrt{\mathcal{C}_A} = 8.86$ pixels. La Figure 2.8 fournit un exemple de résultat de la calibration. Nous avons trouvé des tailles pixel $r_h = 0.174$ mm et $r_v = 0.184$ mm, ce qui est

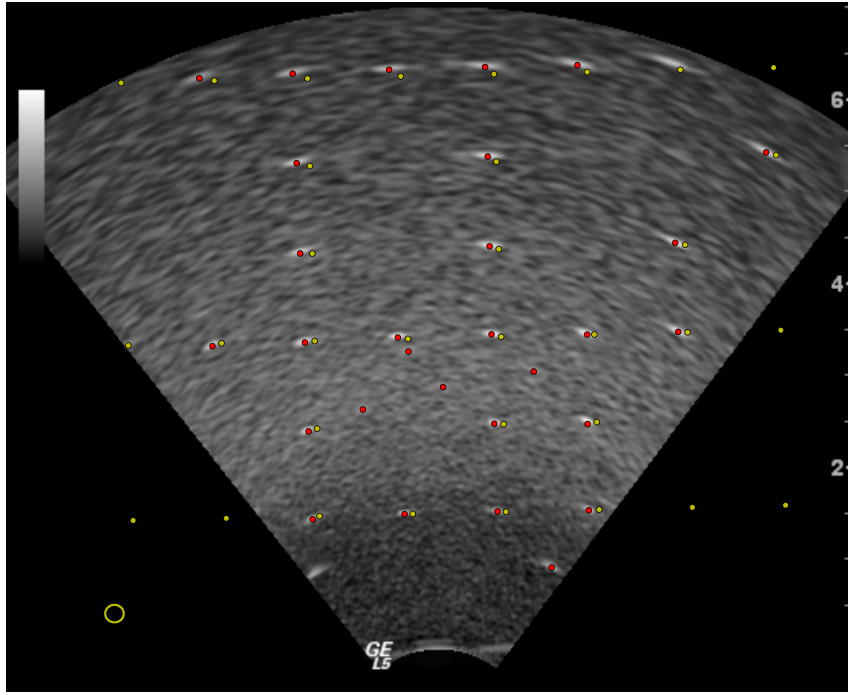


FIGURE 2.8 – Exemple de résultat de la calibration : les points rouges marquent les points cibles automatiquement extraits dans les images, les points jaunes marquent l'intersection du plan image US courant avec le fantôme après recalage. Pour donner une idée du facteur d'échelle, le cercle en bas à gauche a un rayon de 1 mm.

assez différent des données constructeurs (0.1753 mm). Le résidu, exprimé en millimètres, est par conséquent $\sqrt{\mathcal{C}_A} = 1.54$ mm pour la direction horizontale et $\sqrt{\mathcal{C}_A} = 1.63$ mm pour la direction verticale. Ces valeurs sont un peu inférieures au double de l'erreur que présente le système EM, ce qui est à rapprocher de la double utilisation qui en est faite (une fois pour le recalage ${}^{\text{EM}}T_f$, et une autre pour localiser le capteur lié à la sonde).

2.3.4 Synchronisation EM/US/audio

Principe général

La synchronisation des données consiste à mettre temporellement en correspondance toutes les modalités composant le système, qui sont acquises de manière indépendante. Le PC de contrôle, et en particulier son horloge interne, sert de chef d'orchestre : sa précision est de l'ordre de la nanoseconde, ce qui est largement suffisant dans notre système où la fréquence d'acquisition la plus rapide est celle de l'audio (44100Hz, soit une période de l'ordre de $2.27 \cdot 10^{-5}$ s), suivie de la stéréovision à 198Hz. Chaque donnée des modalités acquises par le PC (audio, EM, stéréo) est étiquetée par son temps de réception sur le PC.

Il reste cependant à mesurer le temps écoulé entre l'acquisition de la donnée sur une modalité et sa réception sur le PC de contrôle. Ce temps, appelé délai, peut par exemple correspondre à un temps de traitement interne à la modalité d'acquisition. Il doit être déterminé pour toutes les modalités utilisées.

Afin d'éviter que les erreurs s'ajoutent, nous avons choisi l'audio comme modalité de référence (plus grande fréquence d'acquisition) et avons supposé que le temps entre l'émission d'un son et sa réception par le PC était nul (si on estime à 50 cm la distance entre la bouche du locuteur et le micro, cela correspond à un temps de propagation du son dans l'air de 1.5 ms, soit 7 fois moins que la période des données EM). Le principe général suivi pour la synchronisation a été d'imaginer un dispositif expérimental où un son est émis par un phénomène identifiable dans l'autre modalité (EM, US ou stéréo) : le décalage d'étiquette temporelle entre le son et ce phénomène donne le délai recherché.

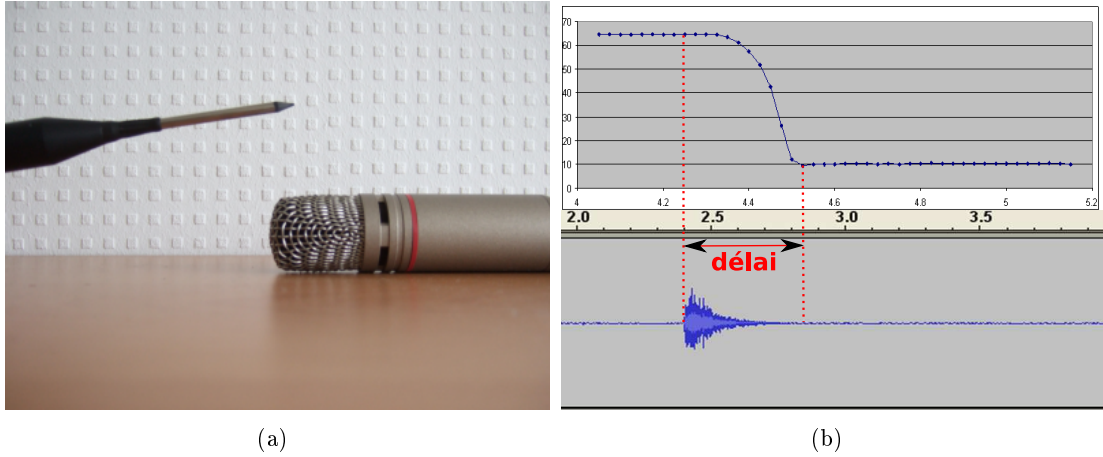


FIGURE 2.9 – (a) Dispositif de synchronisation des données EM avec l’audio. La sonde EM vient taper sur le microphone. (b) En haut, position en Y du stylet EM. En bas, données du fichier audio reçues par le PC de contrôle. Le délai entre les deux modalités est mesuré.

Un premier dispositif expérimental utilisait deux PC [29] : l’un pour enregistrer des données et l’autre pour commander les modalités. Ils étaient synchronisés en leur faisant émettre des bips audio étiquetés par chaque horloge interne, et tous deux capturés sur la piste audio. Cette stratégie s’est simplifiée avec l’achat d’un PC plus puissant permettant de centraliser commande et enregistrement [6].

Synchronisation audio/EM

Le dispositif expérimental déployé pour la synchronisation audio/EM est montré sur la Figure 2.9. Le stylet EM vient frapper le microphone plusieurs fois et les données audio et EM sont enregistrées par le PC de contrôle. Le début du son est identifié dans le fichier audio (Figure 2.9(b), en bas) comme l’est la fin du mouvement du stylet dans les données EM (le générateur de champ magnétique est placé de telle manière que le mouvement du stylet se fasse suivant l’axe Y , voir la Figure 2.9(b), en haut).

Une expérimentation initiale avec la première version du système Aurora à 40Hz a porté sur 20 mesures. En moyenne, le délai constaté a été de 67.8 ms avec un écart-type de 8.9 ms. L’écart-type est donc inférieur à la moitié de la période d’échantillonnage du système EM (25 ms). Cela signifie que le délai entre les deux modalités peut être considéré comme constant.

Cette expérimentation a été reproduite avec le nouveau système fonctionnant à 100Hz. Le nouvel écart trouvé est de 146.2 ms, avec un écart-type de 4.8 ms. Cet écart-type est très légèrement inférieur à la moitié de la période d’échantillonnage des données EM (10 ms). Nous pouvons donc toujours considérer le délai comme constant.

Synchronisation audio/US

L’échographe est un cas particulier car les données ne sont pas acquises sur le PC de contrôle. En revanche, un signal est envoyé par ce PC pour commander à l’échographe de stopper l’acquisition. Les 15 dernières secondes, présentes dans le buffer vidéo, sont alors sauvegardées sur le disque de l’échographe.

Le dispositif expérimental déployé pour la synchronisation audio/US est montré sur la Figure 2.10. Une tige en plastique vient frapper le fond d’un récipient rempli d’eau. La sonde est placée de manière à ce que l’extrémité de la tige apparaisse dans l’image US quand elle vient frapper le fond. Le son produit est capturé par le micro, et étiqueté par le PC.

En frappant plusieurs fois le fond du récipient lors d’une même acquisition US, plusieurs échantillons du délai peuvent être mesurés. Le délai calculé ainsi n’est pas constant (au bruit

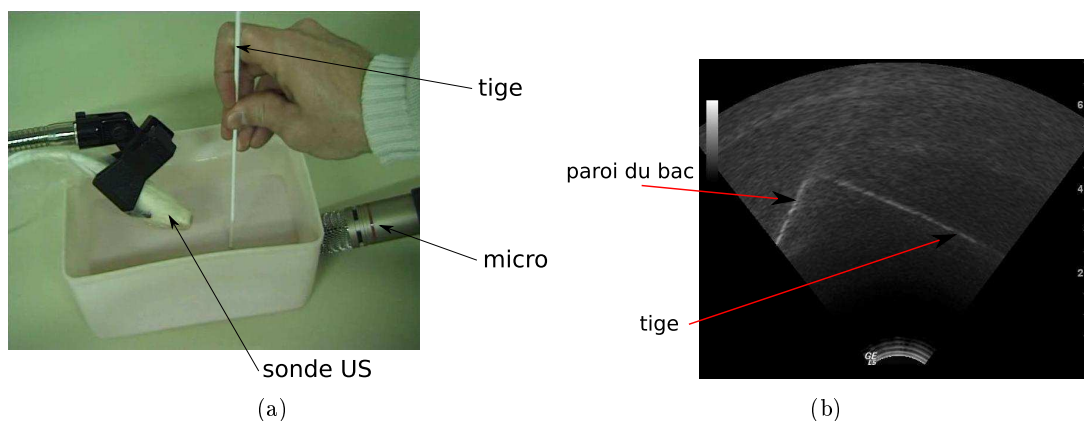


FIGURE 2.10 – (a) Dispositif de synchronisation des données US avec l’audio. Une tige en plastique vient frapper le fond d’un récipient rempli d’eau de telle manière que l’extrémité de la tige apparaît dans l’image US. Le son produit est capturé par le micro. (b) Exemple d’image US à l’instant où la tige frappe le fond du récipient.

de mesure près), mais présente une tendance nette à diminuer au cours de la séquence. La raison en est que la fréquence indiquée par le constructeur pour l’acquisition US n’est pas assez précise. Nous avons donc intégré cette fréquence comme nouvelle variable. Sur 20 expériences, le délai moyen est de 14.9 ms, pour un écart-type de 8.2 ms, ce qui est légèrement supérieur à la demi-période de l’échographe. Nous devons donc intégrer une incertitude de ± 1 image dans l’étiquetage temporel. La fréquence d’acquisition mesurée est de 65.96 Hz (écart-type de 0.02 Hz). La fréquence est donc constante, mais suffisamment différente de celle annoncée pour qu’il soit nécessaire de la calibrer.

2.4 Acquisition et recalage des données IRM

2.4.1 Contexte bibliographique

L’Imagerie par Résonance Magnétique (IRM) a attiré très tôt les chercheurs en étude de la locution. Son premier avantage est que, contrairement à la radiographie, elle ne soumet pas le locuteur à un risque ionisant. L’IRM est une imagerie soit en coupe (en général médio-sagittale pour le conduit vocal, comme l’échographie), soit 3D, qui ne souffre pas des superpositions gênant la lecture d’une radiographie. Son grand avantage sur l’échographie est de produire des images qui couvrent le conduit vocal dans sa quasi-intégralité, même si des structures cristallines comme les dents ou les os ne sont pas visibles. Mais de manière plus fondamentale, l’IRM permet de transposer l’analyse du conduit vocal en 3D, et ouvre la possibilité d’élaboration de modèles plus fins du conduit vocal.

Nous verrons dans les perspectives de ce chapitre que les importants progrès réalisés ces 6 dernières années en acquisition IRM ont fondamentalement changé à notre avis la question de l’acquisition de données articulatoires. Mais à l’époque où ces travaux ont été réalisés, les performances des systèmes IRM imposaient de faire un choix : soit une acquisition rapide (environ 50 Hz) mais uniquement 2D (plan médio-sagittal) et avec une qualité moyenne [SMJC99], soit une acquisition 3D, mais l’état de l’art, qui commençait pourtant à explorer l’acquisition comprimée et l’imagerie parallèle, ne permettait qu’un temps d’acquisition de 10 s pour une résolution de 1.33 mm [KNN09]. Cependant, cette acquisition demandait une machine particulière (3T, rare en 2009) et avec un accès privilégié aux données, ce qui allait à l’encontre de notre objectif de proposer un système aisément reproductible par un autre groupe de recherche.

D’après nos tests, un locuteur moyen peut tenir la phonation d’un phonème pendant environ 20 s. Dès lors, deux stratégies étaient possibles.

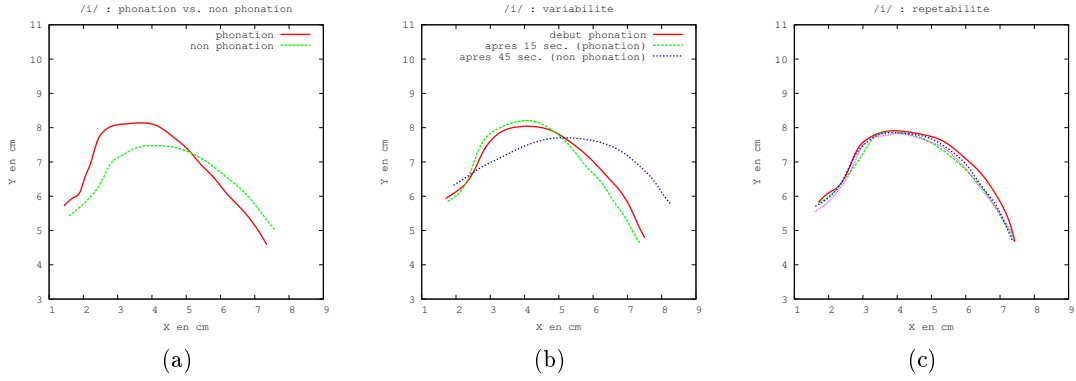


FIGURE 2.11 – Comparaisons de la variabilité et de la répétabilité articutoire pour le phonème /i/. (a) phonation vs. non phonation. (b) variabilité pendant la phonation vs. variabilité entre phonation et non phonation. (c) variabilité due à la répétition.

Protocoles d’acquisition longs

La première stratégie, et la plus couramment utilisée par la communauté parole, consiste à effectuer une seule acquisition *longue* en demandant au sujet de placer ses articulateurs comme s’il prononçait un son, mais sans produire réellement ce son. Badin [BBR⁺02] et Engwall [Eng04] utilisent le mode “spin écho” pour obtenir 53 images sagittales de 256×256 pixels avec une résolution de 1 mm/pixel pour des coupes de 3.6 mm d’épaisseur, espacées tous les 4 mm. L’acquisition d’un phonème dure 43 secondes.

Protocoles avec pauses

Des travaux plus anciens estiment que maintenir la phonation est vital et proposent donc une autre stratégie [BGGN91, STH96] : découper l’acquisition d’un phonème en sous-acquisitions de quelques secondes de phonation entrecoupées de pauses, afin de permettre au sujet de reprendre sa respiration. Story obtient avec un tel protocole en mode “spin écho” 26 coupes sagittales de 256×256 pixels avec une résolution de 0.94 mm/pixel pour des coupes de 5 mm d’épaisseur (l’espacement entre les coupes n’est pas précisé). L’acquisition complète d’un phonème dure 10 min, avec des sous-acquisitions de 8 secondes.

2.4.2 Protocole IRM proposé

Étude préparatoire

Le principal objectif était de proposer un protocole IRM qui ne perturbe pas l’articulation de la parole. Afin de départager les deux stratégies vues ci-dessus, nous avons réalisé une expérience [28], avec pour but d’examiner la variabilité articutoire de phonèmes. Pour ce faire, il est demandé au locuteur d’émettre chaque phonème pendant 15 secondes, puis de stopper la phonation, et de maintenir la position des articulateurs pendant 30 secondes supplémentaires. L’expérience est reproduite plusieurs fois pour chaque phonème. Une image US est capturée en début de phonation ($t=0$ s), une autre en fin de phonation ($t=15$ s) et une dernière en fin d’expérience ($t=45$ s). Les capteurs équipent la sonde et la tête du locuteur afin d’annuler les effets de leurs éventuels mouvements. Le contour est automatiquement segmenté dans chaque image au moyen d’un algorithme que nous avons développé [27].

La Figure 2.11 reproduit les résultats de l’expérience pour le son /i/, qui présente a priori une variabilité moyenne de la position de la langue. Il apparaît clairement que la langue a tendance à revenir à une position neutre, dès que la phonation s’arrête, cette tendance s’accroissant à mesure que le temps sans phonation augmente. Cette tendance a été confirmée chez trois locuteurs, sans

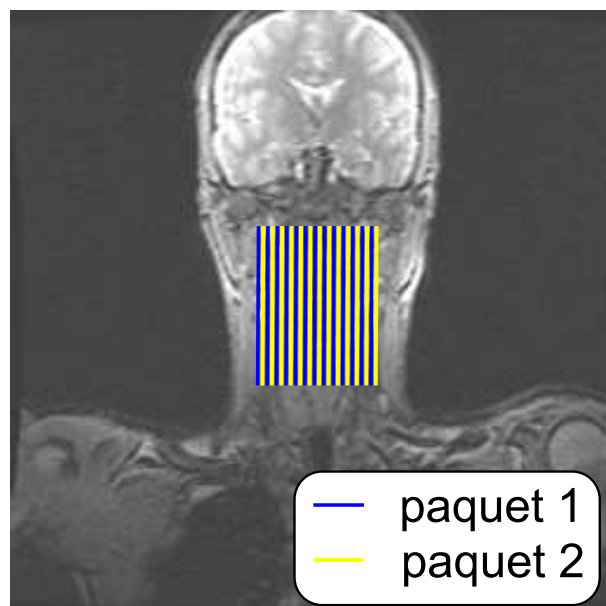


FIGURE 2.12 – Principe d’entrelacement des coupes IRM dans notre protocole. La position des coupes sagittales qui composent une acquisition est indiquée sur fond d’une coupe coronale. Une première série de 13 images est acquise tous les 6 mm (traits bleus), puis une deuxième, identique si ce n’est qu’elle est décalée de 3 mm, est acquise dans un deuxième temps (traits jaunes).

qu’aucun n’ait eu conscience que la position de sa langue n’était pas maintenue. Ce constat disqualifie les protocoles d’acquisition IRM longs.

La répétabilité est en revanche assez bonne : une variation de la position de la langue inférieure à 2 mm est observée à travers trois phonations différentes du /i/. Cette répétabilité est cependant variable selon le phonème et le locuteur : de 2 mm chez un locuteur à 10 mm chez un autre pour les sons /a/ et /u/. Cette variabilité est cependant moindre aux endroits stratégiques pour la phonation (e.g. apex, lieu de constriction maximale). Pour les protocoles avec pauses, il est donc important de ne pas trop prolonger le temps de phonation, et un entraînement du locuteur (sous contrôle échographique par exemple) peut s’avérer bénéfique.

Protocole proposé

Un protocole avec pauses a donc été élaboré avec la collaboration de l’équipe de radiologie du service de neuroradiologie interventionnelle du CHU de Nancy, et sous la supervision du Prof. René Anxionnat. 26 coupes sagittales jointives sont acquises selon un mode T1 Fast Spin Echo (FSE-XL) avec une épaisseur de coupe de 3 mm (une largeur d’environ 8 cm est couverte autour du conduit vocal). L’acquisition se fait en 7 sous-acquisitions de 16 secondes, suivies chacune d’une pause de 10 secondes. Le temps d’acquisition total est donc d’environ 3 minutes par phonème. De plus, les coupes sont entrelacées (13x2) afin de pouvoir aisément détecter un mouvement entre le début et la fin des acquisitions (voir la Figure 2.12). Le tableau 2.1 récapitule les principales caractéristiques du protocole.

Des images IRM dans le plan médiosagittal sont présentées sur la figure 2.13. Ce protocole permet de couvrir l’ensemble du conduit vocal en phonation réelle, avec une résolution dans le plan sagittal divisée par deux par rapport à [BBR⁺02] (images de 512 × 512 pixels à 0.625 mm/pixel au lieu d’images de 256 × 256 pixels à 1 mm/pixel).

Machine	GE Medical Signa HDx 1.5T
Mode	Fast Spin Echo
Nombre de coupes¹	26 (2 paquets de 13)
Épaisseur de coupe	3 mm
Espacement entre les coupes²	2.6 mm et 3.2 mm
TE (echo delay time)	20.712 ms
TR (repetition time)	400 ms
Image	512x512 pixels
Résolution d'image	0.625 mm/pixels

TABLE 2.1 – Résumé des principaux paramètres d'acquisition IRM pour un phonème. Notes : ¹ ce chiffre peut varier suivant la morphologie du sujet. ² le chiffre est donné à titre indicatif, car l'espacement dépend de la position des deux paquets l'un par rapport à l'autre.

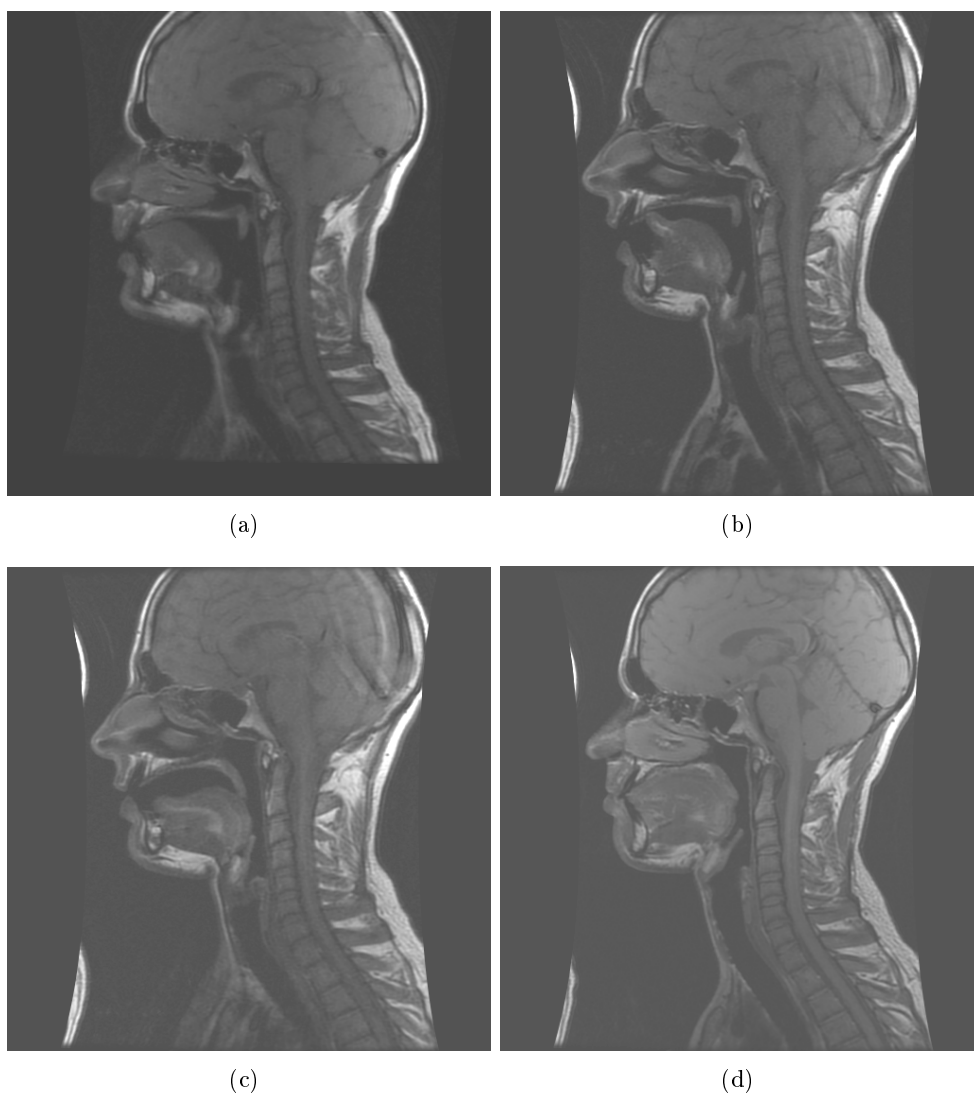


FIGURE 2.13 – Exemples d'images IRM acquises (coupes médiosagittales) : (a) /i/ (b) /l/ au début de l'articulation du /la/ (c) /â/ (d) IRM de référence, position au repos.

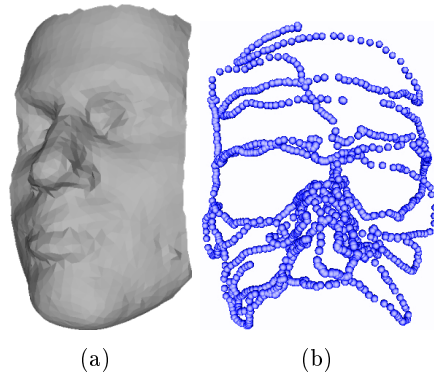


FIGURE 2.14 – Exemple de données utilisées pour recalage de l’IRM avec les capteurs EM : (a) la surface du visage extraite par seuillage de l’IRM de référence, (b) le nuage de points 3D obtenus par balayage du visage du locuteur avec le stylet EM.

2.4.3 Recalage des images IRM

Entre deux sessions, voire entre deux acquisitions lorsque la session est longue, la tête du locuteur peut bouger dans l’IRM⁸. Il est donc nécessaire de pouvoir recalrer les différentes IRM entre elles. Par ailleurs, l’intégration de ces données IRM dans le système ARTIS requiert qu’elles soient spatialement alignées avec le repère de la tête du locuteur, qui sert de référence dans le système.

Problème monomodal

Lors de chaque session d’acquisition IRM, nous avons pris soin d’acquérir une image 3D du locuteur au repos : le visage et les articulateurs sont en position neutre. Cette image est également acquise avec un protocole T1 Fast Spin Echo, mais ici classique, sur un temps de 4 minutes.

Chaque acquisition de phonème est ensuite recalée en 3D en maximisant l’information mutuelle avec l’image de référence [PMV03]. Les critères iconiques ont en effet été reconnus comme plus précis que les critères structurels (e.g. basés contours ou points) à condition de disposer d’une bonne initialisation pour le recalage [MHH⁺05], ce qui est notre cas puisque les mouvements de la tête sont de faible amplitude. Par ailleurs, un des problèmes majeurs est que le recalage ne soit pas influencé par la position variable des articulateurs, notamment la mâchoire inférieure, et puisse se concentrer sur la partie haute du crâne qui définit une référence stable. Le critère d’information mutuelle à optimiser est un critère statistique robuste aux données aberrantes et de ce fait constitue un excellent choix. Nous avons employé l’algorithme proposé par Mattes [MHV⁺03], et implémenté dans ITK.

Problème multimodal

L’échographie est une modalité 2D en coupe, qui par ailleurs ne montre essentiellement que la langue, articulateur très mobile. Les données EM sont 3D, mais il est impossible d’utiliser les capteurs métalliques dans l’IRM. Au début de chaque acquisition avec le système ARTIS, nous avons donc numérisé le visage du locuteur, en position neutre, en balayant sa surface avec le stylet EM. Les données ainsi acquises sont exprimées dans le repère de référence \mathcal{R}_0 . Par ailleurs, la surface du visage est extraite par simple seuillage dans l’IRM de référence. La Figure 2.14 donne un exemple des données utilisées pour ce recalage. Le nuage de points EM est aligné sur la surface IRM par un algorithme ICP classique (Iterative Closest Point) [BM92].

8. Nous rappelons qu’un bouger lors d’une acquisition fractionnée est détecté grâce à l’entrelacement, auquel cas le phonème est de nouveau acquis.

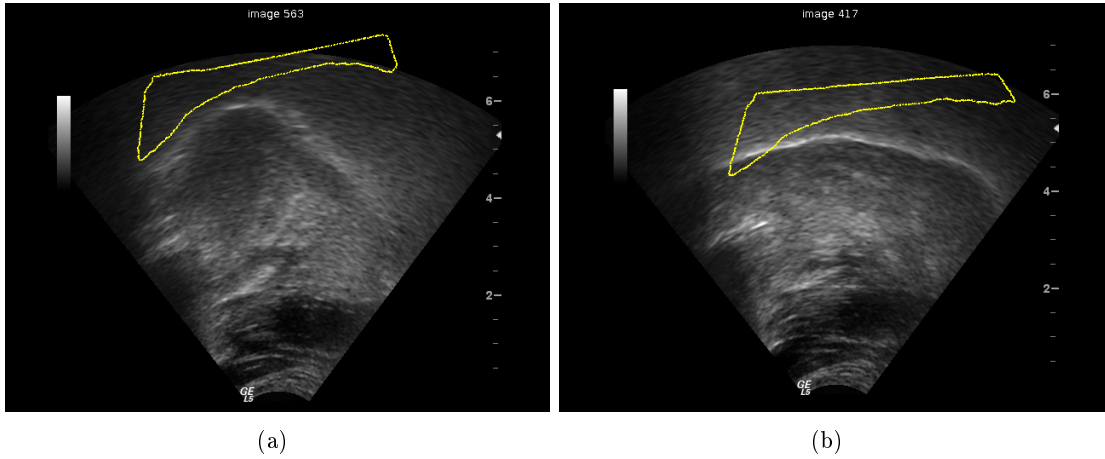


FIGURE 2.15 – Exemples de contours de palais extraits de données IRM et superposés après recalage dans les images US : (a) /u/ de /au/, (b) (a) transition /g/ → /e/ de «nager».

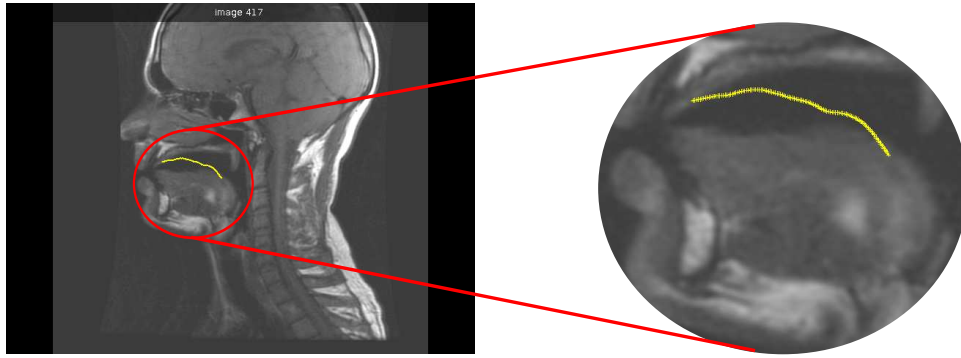


FIGURE 2.16 – Visualisation du contour de langue extrait de l'US dans une image IRM pour la transition /g/ → /e/ de «nager».

2.5 Quelle fiabilité pour ces données ?

L'étude reprise dans cette section n'a pas fait l'objet d'une publication, si ce n'est dans le manuscrit de thèse de Michaël Aron [Aro09]. Nous n'en reprenons que les points essentiels.

2.5.1 Évaluation visuelle

Le système ARTIS intègre un certain nombre de modalités d'acquisition différentes, à la fois spatiales et temporelles. Nous avons vu que des protocoles de synchronisation simples à mettre en œuvre permettent d'obtenir un excellent alignement temporel des modalités (1 image US de décalage dans le pire cas). En revanche, l'incertitude sur les données spatiales mérite une attention particulière. Pour s'en convaincre, nous reprenons sur la Figure 2.15 un exemple de superposition du palais, dont la surface est extraite par seuillage de l'IRM, sur des images US. La superposition est cohérente pour la majorité des séquences (par exemple le /u/ de /au/ sur la Figure 2.15(a)) mais semble échouer sur d'autres (par exemple la transition entre le /g/ et le /e/ de «nager» sur la Figure 2.15(b)), sans qu'une cause quelconque puisse le justifier. Cette mauvaise impression visuelle est toutefois à relativiser. Un autre choix iconographique suffit à faire disparaître cette incohérence : en choisissant comme image de fond l'IRM et en y superposant le contour de la langue extrait de l'image US, pour le même phonème que la Figure 2.15(b), la relation spatiale entre la langue et la palais semble tout à fait valide (voir la Figure 2.16). Ceci se justifie très aisément par le rapport des tailles pixel US (environ 0.17 mm) et IRM (0.625 mm), qui implique qu'un seul pixel en IRM couvre environ 16 pixels dans l'image US.

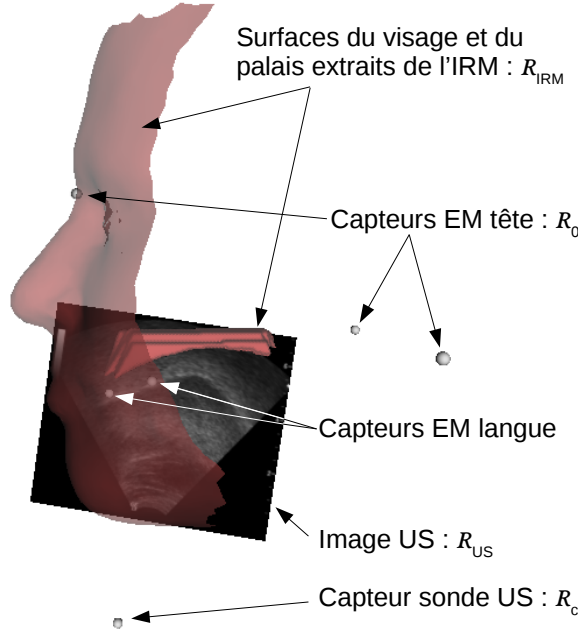


FIGURE 2.17 – Récapitulatif des données et repères du système multimodal ARTIS.

2.5.2 Récapitulatif des repères et transformations

Une analyse plus quantitative nécessite de récapituler les repères et transformations en jeu. Dans la lignée de ce texte, nous excluons à nouveau la stéréovision, sans que cela remette en cause le principe général suivi pour cette étude.

Trois modalités sont impliquées :

- L'échographie (US), dont les données sont exprimées dans le repère \mathcal{R}_{US} . Ce repère recouvre en fait deux repères : l'un dont l'unité est le pixel et l'autre le millimètre. La transformation pour passer de l'un à l'autre est l'homothétie centrale notée par S_{US} paramétrée par les tailles pixel US r_h et r_v .
- L'IRM, dont les données sont exprimées dans le repère de l'IRM de référence \mathcal{R}_{IRM} . Ici, nous ne considérons que le repère en millimètre, la taille voxel IRM étant supposée calibrée avec précision.
- Les données EM qui recouvrent en fait trois repères millimétriques : le repère \mathcal{R}_{EM} lié au générateur de champ magnétique (c'est dans ce repère que sont localisés par défaut les capteurs EM), le repère de référence \mathcal{R}_0 défini par les trois capteurs placés sur la tête, et donc solidaire de celle-ci, et le repère \mathcal{R}_c lié au capteur solidaire de la sonde US.

La Figure 2.17 reprend les différents composants du système multimodal ARTIS, chacun ayant un repère associé.

Outre S_{US} , les autres transformations élémentaires permettant de changer de repère sont les transformations rigides :

- ${}^US T_c$: calibrée grâce à notre fantôme, sa détermination fait intervenir des transformations intermédiaires en lien avec \mathcal{R}_f (voir la Section 2.3.2) ;
- ${}^{IRM} T_{EM}$: calibrée par recalage de la surface du visage par ICP ;
- ${}^0 T_{EM}$: déduite des données des trois capteurs 5DDL fixés sur la tête du locuteur ;
- ${}^c T_{EM}$: lue sur le capteur solidaire de la sonde US.

Les autres transformations sont déduites de ces transformations élémentaires par combinaisons successives. Ainsi la transformation qui nous a intéressé plus particulièrement, et qui implique la plus longue chaîne de combinaisons, est celle qui relie l'US à l'IRM :

$${}^{IRM} T_{US} = {}^{IRM} T_{EM} \cdot {}^{EM} T_c \cdot {}^c T_{US} \cdot S_{US} \quad (2.6)$$

Combinaison à laquelle il faudra intégrer ${}^0 T_{EM}$ si l'on veut obtenir toutes les données dans \mathcal{R}_0 .

2.5.3 Évaluation quantitative

Au sein du système ARTIS, le recalage spatial s'implémente à travers la détermination de différents changements de repère, transformations dont le calibrage a été au cœur de notre apport à ce projet. Cependant, chaque transformation est calculée selon un processus plus ou moins complexe à partir de données qui sont bruitées. Si le niveau de bruit est acceptable pour l'étude de la parole selon chaque modalité, il est vital de quantifier le niveau de bruit que présente le système multimodal : soit l'incertitude est suffisamment faible pour être négligée (selon l'avis des spécialistes en parole et l'application visée), soit elle est plus importante et il est essentiel de la connaître afin de la prendre en compte pour interpréter correctement les résultats.

Méthode suivie

La manière la plus fiable d'estimer l'incertitude d'un recalage, consiste à calculer l'erreur de recalage sur une cible (TRE pour *Target Registration Error*) [FWMJ98, Fit09]. Cela nécessite une vérité terrain dont nous ne disposons pas pour le système ARTIS. Une alternative consiste à employer une technique de propagation d'incertitude : la covariance du bruit sur les données en entrée, peut être propagée analytiquement aux transformations et leurs combinaisons [PT97, NPSA04]. Elle repose cependant sur l'hypothèse que le gradient du critère à minimiser est suffisamment régulier, notamment vis-à-vis des données d'entrée. Ce n'est pas notre cas où certaines transformations (${}^{\text{US}}T_c$ ou ${}^{\text{IRM}}T_{\text{EM}}$ par exemple) sont calculées par un processus fortement non-linéaire. Nous avons donc choisi de procéder par simulation de Monte-Carlo [HZ00].

Incertitude sur les données

Les incertitudes admises en entrée de ces simulations ont été déterminées comme suit. L'erreur sur les données EM a été mesurée expérimentalement. Un fantôme dédié à la calibration de la taille pixel US (Modèle 555, CIRS Inv., Norfolk, VA), semblable au nôtre, a conduit à une estimation de l'incertitude de 0.007 mm/pixel pour les tailles pixel r_h et r_v . Nous avons assimilé l'incertitude de la localisation des données IRM à la taille voxel (0.625mm×0.625mm×3mm). Enfin, pour les images US, le bruit important, et des questionnements sur la localisation exacte des interfaces dans les images, nous ont incité à évaluer la précision d'un pointage manuel (lignes du fantôme selon diverses incidences, par trois observateurs). L'incertitude a ainsi été évaluée 2.09 pixels. La distribution sur les données EM et US sont supposées suivre une loi gaussienne et celles sur les données IRM et la taille pixel US suivent une loi uniforme.

Expérimentation

Nous avons étudié la transformation ${}^{\text{MR}}T_{\text{US}}$ donnée par l'équation 2.6, a priori la plus sensible à l'incertitude du système. Quatre points 2D correspondant à des positions de la langue et du palais dans les images US, ont été choisis comme points de mesure. La variance sur leur position dans le repère IRM à la suite de 1000 simulations de Monte Carlo a été mesurée entre 2.32 mm et 2.53 mm selon le point.

Ce niveau d'incertitude correspond à 15 pixels dans l'image US (4 en IRM, voire moins de 1 dans la direction orthogonale au plan sagittal). Si on tient compte de cette erreur dans la position du palais superposé à l'image US, de nombreuses interrogations sont levées. Ainsi pour une séquence où 38% des images posaient question (par exemple, la langue semblait traverser le palais), le nombre de tels cas a chuté à 4% en tenant compte d'une erreur de localisation du palais possible de 15 pixels. La Figure 2.18 superpose à la Figure 2.15 un cercle qui matérialise l'incertitude sur la position du palais : il existe une position dans ce cercle pour laquelle la langue ne traverse plus le palais.

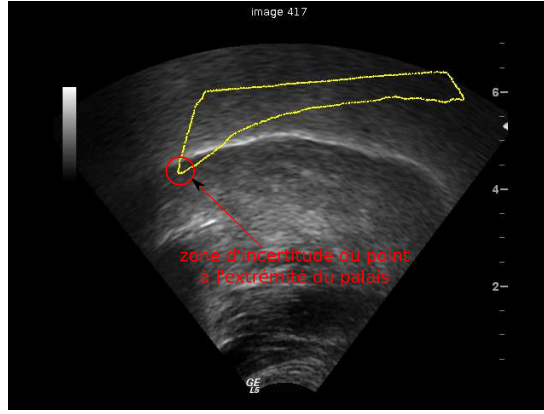


FIGURE 2.18 – Superposition d’un cercle de 15 pixels pour matérialiser le niveau d’incertitude dans une fusion US/IRM : même si la langue semblait traverser le palais dans l’image originale, il existe une position à l’intérieur du cercle pour laquelle le palais est simplement en contact avec la langue.

2.6 Exemple d’exploitation du système ARTIS

Le système ARTIS a été exploité, dès le début, par des chercheurs partenaires du projet, spécialistes de l’étude de la parole. Nous reprenons ici une étude publiée dans [25], comme exemple de l’intérêt que présente un système d’acquisition de données articulatoires multimodales.

L’objectif de cette étude était d’évaluer la cohérence des données acquises avec le modèle articulatoire statistique proposé par Maeda [Mae79], et largement utilisé par la communauté parole. Ce modèle décrit la forme du conduit vocal par le profil de sa largeur, échantillonné le long d’une grille semi-polaire de segments superposés aux contours du conduit vocal dans une coupe médio-sagittale (Figure 2.19). Toutes les données de largeur de conduit vocal sont centrées et normalisées. Ainsi, pour un son donné, la forme du conduit vocal est décrite par un vecteur de 29 données v que Maeda a relié linéairement à 7 paramètres articulatoires p par une Analyse en Composantes Principales :

$$v = Ap \quad (2.7)$$

où A est une matrice constante qui ne dépend que du locuteur. Le système ARTIS n’offrant aucune donnée sur le larynx, nous n’avons considéré que les 6 premiers paramètres articulatoires. A est donc dans notre cas une matrice 29×6 .

Le système ARTIS permet de localiser dans un même repère, le contour de la langue, segmenté dans les images US, avec celui du palais, segmenté en IRM. Il est donc possible de mesurer partiellement le profil de largeur du conduit vocal, et d’obtenir un sous-système de l’équation 2.7. Les paramètres articulatoires p peuvent alors se recouvrer par inversion.

Le modèle de Maeda est en particulier bien adapté aux voyelles. Nous avons donc acquis des transitions voyelle-voyelle (/ae/ et /ay/ sont données en exemple ici), ce qui a permis de compenser la partialité de nos données par une régularisation sur p à deux niveaux : d’une part, chaque paramètre articulatoire étant normalisé, nous avons imposé la contrainte $p_i \in [-3, 3]$ pour restreindre la variation de chaque paramètre dans la limite de 3 écarts-types ; d’autre part, un terme de régularisation temporelle, impliquant vitesse et accélération, a été ajouté pour la minimisation du critère quadratique d’attache aux données [25].

La cohérence des données ARTIS avec le modèle de Maeda est évaluée en superposant les contours de la langue extraits des images US, sur la forme du conduit vocal reconstruite à partir des paramètres articulatoires inversés. La Figure 2.20 reprend cette comparaison lors de la transition VV /ae/. Les contours reconstruits sont proches des contours US sur la partie qui leur est commune. D’un point de vue plus articulatoire, la forme reconstruite est celle à laquelle on peut s’attendre pour cette articulation, en particulier les lieux de constriction sont bien prédits.

La Figure 2.21 montre que le modèle a plus de difficulté à s’ajuster à nos données pour la transition VV /ay/. Ce n’est pas un problème de recalage car le décalage n’est pas constant.

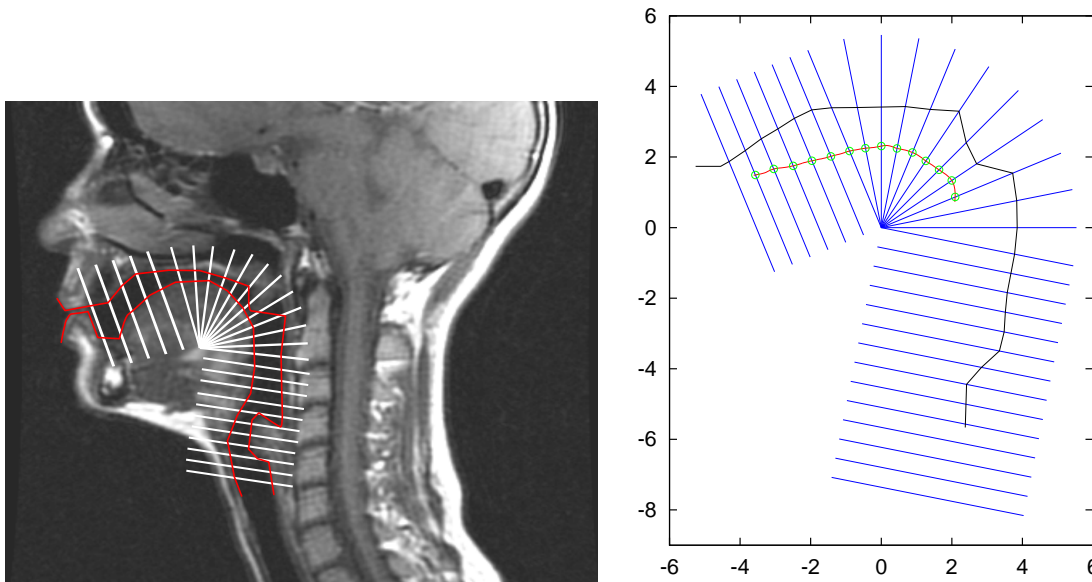


FIGURE 2.19 – Modèle de Maeda : (gauche) Grille semi-polaire du modèle de Maeda (en blanc) surposée aux contours du conduit vocal (en rouge) dans une coupe médio-sagittale en IRM ; (droite) superposition de la grille sur des contours de la langue extraits d’images US. L’intersection avec la grille ne fournit que des données partielles.

Une erreur peut arriver même lorsque le contour de la langue dans les images US est clair et univoque. Il s’agit donc probablement d’une limite de ce modèle. Cela nous a amené à reconsidérer la construction du modèle articulatoire.

2.7 Recalage et segmentation automatique d’images IRM

Le modèle de Maeda a été construit pour une locutrice spécifique, et pour l’étude qui précède, nous l’avons adapté à notre locuteur en appliquant une similitude, comme cela est fait classiquement. Ce procédé est sans doute trop simple. Nous avons donc envisagé la question de construire un modèle multi-locuteurs. Deux questions se sont posées. D’une part, le modèle de Maeda se base sur des séquences cinéradiographiques, qu’il n’est plus possible d’acquérir aujourd’hui. Par ailleurs, les contours y ont été délimités à la main, et leur qualité souffre donc, outre des ambiguïtés inhérentes aux projections rayons X, de la subjectivité des manipulateurs. Nous avons donc mis au point une méthode de segmentation de la langue dans les images IRM en coupe médio-sagittale, adaptée à la construction de modèle articulatoire statistique multi-locuteurs. Ce travail a fait l’objet du stage post-doctoral de Ting Peng en 2009. Notre approche a été décrite dans [23].

2.7.1 Hypothèses et données

Les données IRM sont les mêmes que pour notre protocole décrit en Section 2.4 : pour chaque locuteur, nous disposons d’une acquisition 3D, dite de référence, en haute résolution de la tête en position neutre, et d’une acquisition rapide, selon notre protocole, pour chaque phonème ; chaque IRM de phonème est recalée sur l’IRM de référence par maximisation de l’information mutuelle.

Si tous les recalages sont effectués en 3D, la segmentation n’a lieu quand dans la coupe identifiée comme médio-sagittale dans l’IRM de référence, ce qui implique donc un reformattage de cette coupe pour les IRM de phonèmes.

Nous choisissons parmi les locuteurs un locuteur dit de référence. La langue est segmentée à la main par plusieurs manipulateurs, dans les IRM de phonèmes, et un modèle articulatoire est construit pour ce locuteur par ACP [CTCG95]. Ce locuteur donc sera choisi dans la base de données des locuteurs, parmi ceux pour lesquels nous disposons d’un plus grand nombre

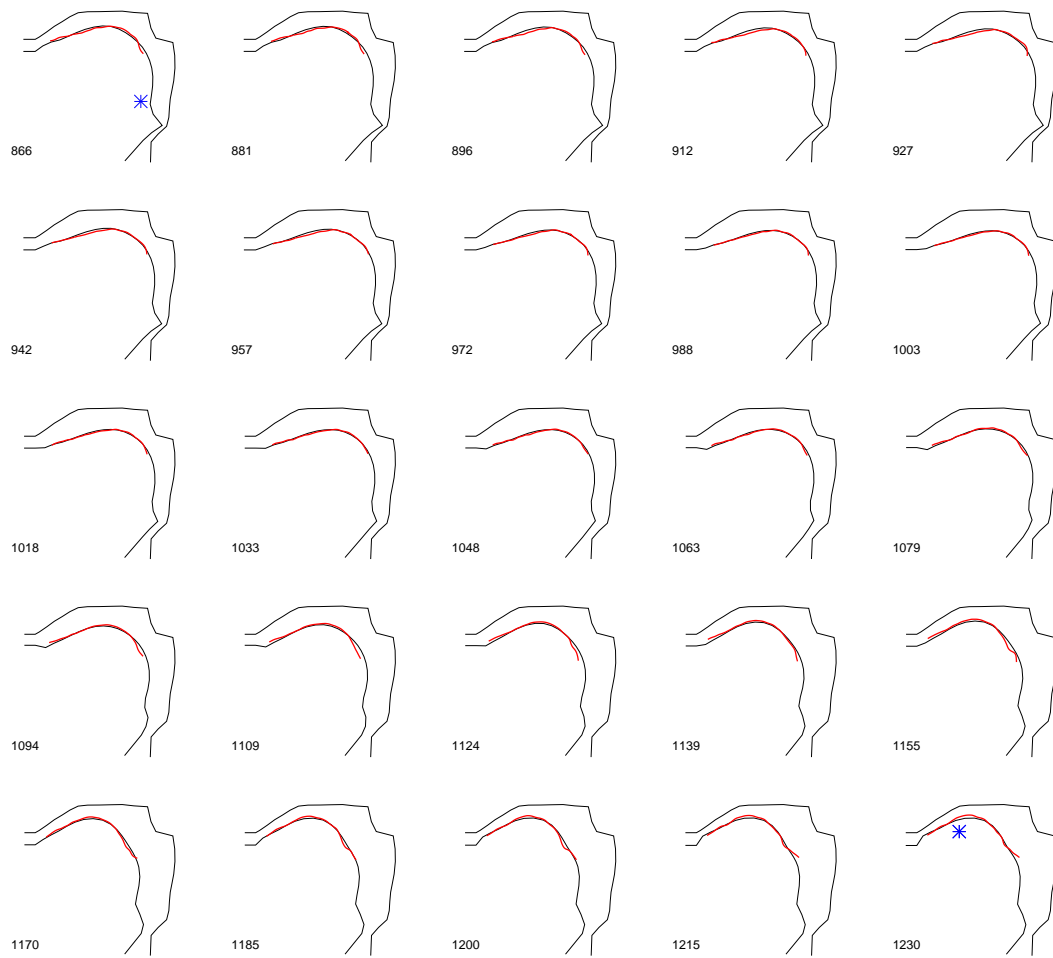


FIGURE 2.20 – Contours US superposés au modèle de Maeda ajusté à ces données pour la séquence VV /ae/. Le temps en millisecondes est indiqué sur chaque forme. Les astérisques indiquent le lieu de constriction.

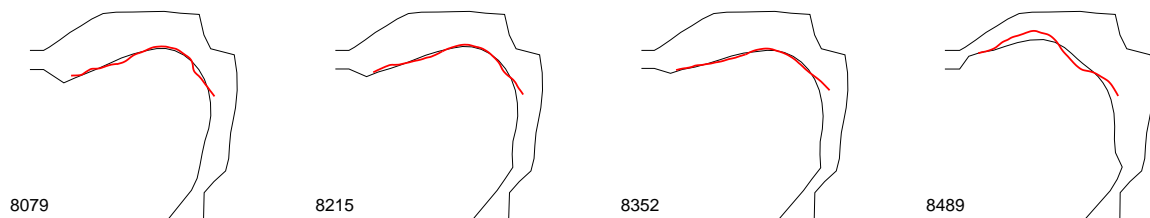


FIGURE 2.21 – Contours US superposés au modèle de Maeda ajusté à ces données pour la séquence VV /ay/. Le temps en millisecondes est indiqué sur chaque forme.

de phonèmes imagés, afin d’assurer autant que faire se peut la bonne couverture du domaine articulatoire par l’ACP.

2.7.2 Méthode

Recalage inter-locuteurs

Une grande diversité morphologique existe chez les locuteurs. Afin de la compenser, nous avons appliqué un recalage affine, basé sur la maximisation de l’information mutuelle. Ce recalage a été effectué en 2D sur une IRM de phonation de voyelle où le conduit vocal est ouvert afin d’avoir un bon ancrage du recalage sur le palais. Ce recalage grossier est appliqué aux autres IRM phonèmes. Il permet de corriger les différences de taille et de position entre locuteurs, mais les stratégies d’articulation différentes selon les locuteurs peuvent entraîner un décalage résiduel (r_1, r_2) en translation.

Modèle de courbe

Une des critiques qui peut être faite au modèle de Maeda émane d’un changement de point vue. Au lieu de considérer la largeur du conduit vocal, considérons la position du point d’intersection de la langue avec chaque segment de la grille semi-polaire. Ce point de vue est équivalent au précédent puisque l’autre point d’intersection, avec le palais, est fixe. Lors du mouvement de phonation, les points identifiés sur un même segment à la surface de la langue ne se correspondent pas physiquement. L’ACP déduite sous-estime par conséquent les mouvements de langue parallèles à la ligne centrale du conduit vocal. L’effet de ce manque de correspondance physique a été très bien documenté [DTC⁺02] et certaines solutions d’appariement automatique ont été proposées dans le cas de courbes fermées en se basant sur la forme des courbes, par exemple en supposant une similarité de courbure [Tag99], ou bien en s’intéressant plutôt au modèle paramétrique estimé, par exemple en cherchant à minimiser la longueur de description du modèle (*Minimum Description Length*) [DTC⁺02].

Dans le cas d’objets fortement déformables et sans caractéristique de forme aisément exploitable, comme c’est le cas en général pour les organes, ces critères échouent souvent et nous avons proposé dans [8] une méthode semi-manuelle où le contour fermé est segmenté manuellement en sections ayant un sens, par exemple anatomique, pour l’expert. Les extrémités de ces sections sont mise en correspondance, et la correspondance est étendue par un appariement basé sur l’abscisse curviligne par rapport à chaque section. Cette méthode s’étend naturellement à une courbe ouverte, à condition d’en identifier les extrémités de façon univoque. Nous avons localisé le contour de la langue entre la base du frein lingual et celle de l’épiglotte (voir la Figure 2.22). Nous avons gardé les $p = 15$ composantes principales δC_i afin de capturer 99.87% de la variance. Un a priori de forme est ainsi introduit dans le processus de segmentation en imposant, à travers leur paramétrisation, que les courbes fassent partie de l’espace à 15 dimensions ainsi défini.

En conséquence, les courbes C recherchées sont définies par 17 paramètres (w, r) :

$$C(s; w, r) = \bar{C}(s) + \sum_{i=1}^p w_i \delta C_i(s) + \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (2.8)$$

où s est l’abscisse curviligne, \bar{C} est la courbe moyenne, et les $\{w_i\}_{i \in \llbracket 1, p \rrbracket}$ sont les coefficients propres de la courbe dans la base des composantes principales. Cette formulation est à rapprocher de [TYWI⁺03] où l’ACP modélise cependant la fonction caractéristique de la courbe, et non la courbe elle-même.

En pratique, les courbes sont échantillonnées sur N points $P_i = (x_i, y_i)$, et sont donc représentées comme des vecteurs à $2N$ dimensions $C = (x_1, y_1, x_2, y_2, \dots, x_N, y_N)$, en accord avec le formalisme employé pour l’ACP. L’équation précédente se réécrit alors :

$$C(w, r) = \bar{C} + \sum_{i=1}^p w_i \delta C_i + r_1 \delta r_1 + r_2 \delta r_2 \quad (2.9)$$

où $\delta r_1 = (1, 0, 1, 0, \dots, 1, 0)$ et $\delta r_2 = (0, 1, 0, 1, \dots, 0, 1)$.

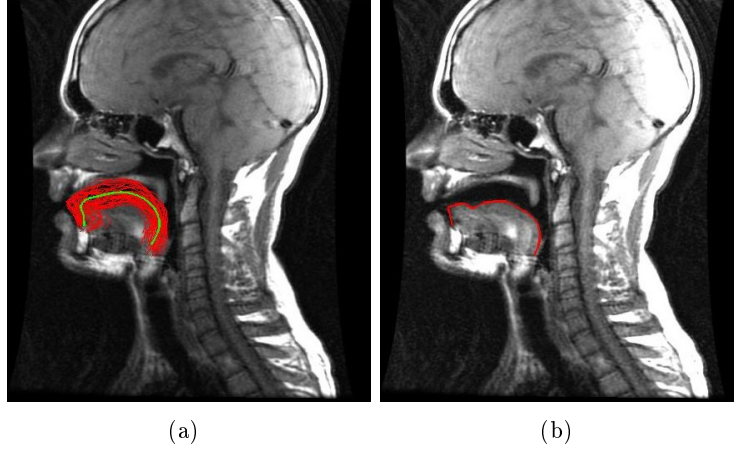


FIGURE 2.22 – (a) Contours utilisés pour construire l’ACP chez le locuteur de référence, avec la courbe moyenne en vert ; (b) Les extrémités de chaque contour sont fixées à la base du frein lingual (à gauche) jusqu’à celle de l’épiglotte (à droite)

Énergie

Nous proposons d’évaluer la qualité de la segmentation de l’image I proposée par la courbe C , en calculant l’énergie :

$$E(C; I) = \alpha E_G(C; I) + E_L(C; I) \quad (2.10)$$

avec

$$E_G(C; I) = \int_{C_{\text{in}}} (I(P) - \mu)^2 dP + \beta \int_{C_{\text{out}}} (I(P) - \nu)^2 dP \quad (2.11)$$

et

$$E_L(C; I) = \int_{\Omega} \left\{ \int_{C_{\text{in}}} K(P - Q) (I(Q) - u(P))^2 dQ + \int_{C_{\text{out}}} K(P - Q) (I(Q) - v(P))^2 dQ \right\} dP \quad (2.12)$$

Dans ces équations, Ω désigne le support de l’image, C_{in} et C_{out} respectivement l’intérieur et l’extérieur de la courbe C .

E_G est la fonctionnelle proposée par Chan et Vese [CV01] qui fait l’hypothèse que la courbe C est le contour d’un objet uniforme, d’intensité moyenne μ , disposé sur un fond uniforme, d’intensité moyenne ν dans l’image. La mise à jour de μ et ν , ainsi que le calcul intégral dans E_G nécessitent que C soit fermée, ce que nous réalisons en reliant ses deux extrémités par un segment. Cette énergie permet de positionner la courbe de manière robuste dans l’image. La langue étant un muscle, et un des seuls visibles dans une image d’IRM du crâne, l’hypothèse d’un objet uniforme est valide. Mais ce n’est pas le cas pour le fond, beaucoup plus varié.

E_L permet de compenser les erreurs induites par les perturbations du fond en calculant la même énergie, mais restreinte, en chaque point de mesure P , à son voisinage [LKGD07]. K est un noyau, en pratique gaussien, qui permet de définir un voisinage par pondération des pixels Q en fonction de leur distance à P . D’une manière similaire à E_G , $u(P)$ et $v(P)$ sont les moyennes des intensités intérieure et extérieure, mais calculées ici localement en employant la même pondération par K (voir [LKGD07] pour plus de détails).

Les poids α et β permettent d’équilibrer les influences des termes d’énergie (globale vs locale, intérieur vs extérieur). Il est intéressant de noter que, contrairement à d’autres formulations de contours actifs (e.g. snakes), nous n’avons pas besoin de régularisation par un terme d’énergie interne (e.g. longueur et/ou courbure), car celle-ci est assurée par l’a priori de forme fourni par l’ACP.

Optimisation

Nous avons minimisé E , par rapport aux 17 paramètres $\{w_i\}$ et (r_1, r_2) , par une descente de gradient. Les équations d'évolution s'obtiennent alors par dérivation en chaîne :

$$\frac{\partial w_i}{\partial t} = -\frac{\partial E}{\partial w_i} = -\frac{\partial E}{\partial C} \frac{\partial C}{\partial w_i} = -\frac{\partial E}{\partial C} \delta C_i \quad (2.13)$$

$$\frac{\partial r_i}{\partial t} = -\frac{\partial E}{\partial r_i} = -\frac{\partial E}{\partial C} \frac{\partial C}{\partial r_i} = -\frac{\partial E}{\partial C} \delta r_i \quad (2.14)$$

La dérivée de l'énergie par rapport à la courbe en un point x de celle-ci se déduit des équations d'évolution données par [CV01, LKGD07]

$$\begin{aligned} \frac{\partial E}{\partial C(x)} = & \left\{ \alpha \{ (I(x) - \mu)^2 - \beta (I(x) - \nu)^2 \} \right. \\ & \left. + \int_{\Omega} K(y - x) (I(x) - u(y))^2 dy - \beta \int_{\Omega} K(y - x) (I(x) - v(y))^2 dy \right\} \mathbf{N}(x) \end{aligned} \quad (2.15)$$

où $\mathbf{N}(x)$ est la normale extérieure à la courbe en x [TYW01].

Le vecteur $\frac{\partial E}{\partial C}$ est alors formé en rassemblant les 2-vecteurs $\frac{\partial E}{\partial C(x)}$ calculés en chaque point échantillonné sur la courbe de manière similaire aux δC_i . Les courbes sont restreintes à des formes raisonnables en contraignant chaque poids w_i associé à une composante principale à rester dans l'intervalle $[-5\sqrt{\lambda_i}, 5\sqrt{\lambda_i}]$ où λ_i est la variance attachée à cette composante principale dans l'ACP.

2.7.3 Résultats

Nous avons évalué cette procédure de segmentation sur trois locuteurs – deux hommes H_1 (34 articulations) et H_2 (20 articulations), et une femme F_1 (10 articulations) – en prenant comme référence l'ACP réalisée à partir de la segmentation manuelle d'IRM d'un homme H_0 (39 articulations). La courbe est toujours initialisée par le contour dessiné chez H_0 pour l'articulation étudiée. Le paramètre β , entre intérieur et extérieur de la courbe, a été fixé à 1 dans la plupart des expériences, sauf quand le contour initial était trop éloigné du contour recherché. Le poids α , entre énergie globale et locale, a été adapté au locuteur ($\alpha = 0.08$ pour H_1 et H_2 , et $\alpha = 0.045$ pour F_1).

La Figure 2.23 illustre la qualité des contours obtenus (voir [23] pour plus de résultats). Le processus a cependant subi quelques rares échecs dont nous donnons deux exemples sur la Figure 2.24 : le contour de la langue n'est pas clairement défini dans l'image de la Figure 2.24(a), conséquence d'un effet de volume partiel ; dans la Figure 2.24(b), l'apex est très proche de la base du frein lingal et accroche donc l'extrémité du contour.

Enfin, l'importance de combiner les deux énergies globale de Chan et Vese [CV01], et locale de Li et al. [LKGD07] apparaît dans la Figure 2.25

2.8 Conclusion et Perspectives

Dans ce chapitre, nous avons passé en revue les résultats de nos travaux dans le domaine de l'acquisition de données articulatoires. Bien qu'à la marge de notre domaine applicatif de prédilection, ce projet présente toutes les caractéristiques d'un projet en imagerie médicale interventionnelle et multimodale :

- un besoin d'observation de structures anatomiques internes, en mouvement et fortement déformables ;
- l'absence d'une modalité d'acquisition permettant à elle seule de couvrir l'ensemble des besoins en observation ;
- le recours à des modalités d'imagerie médicale ou de capture de mouvement également utilisées dans un cadre clinique ;

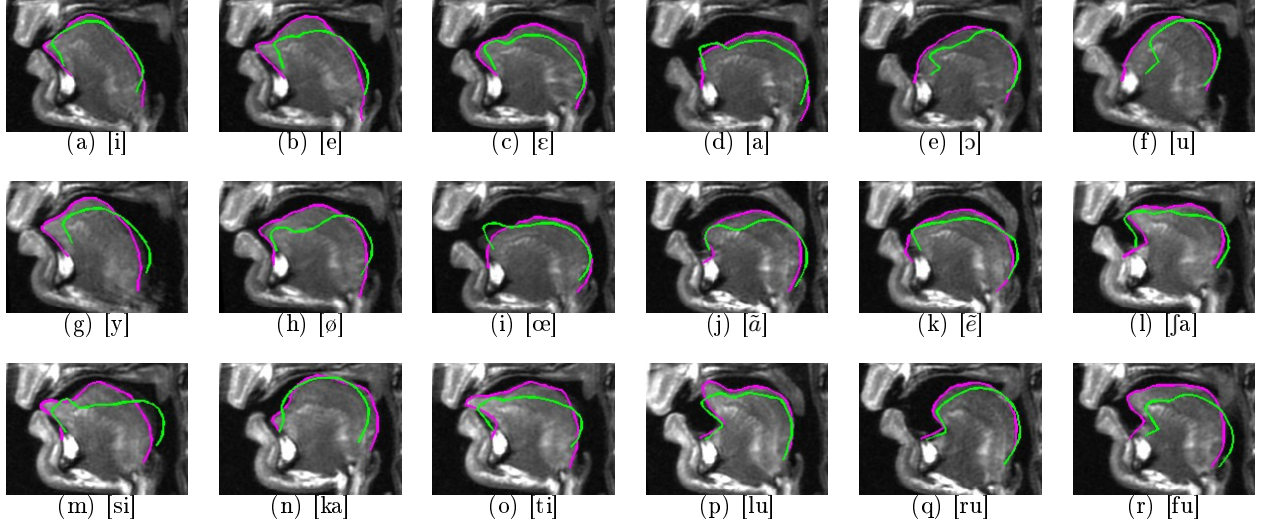


FIGURE 2.23 – Résultats pour H_1 . La courbe verte donne le contour initial, et la courbe magenta la courbe finale obtenue par notre procédé de segmentation.

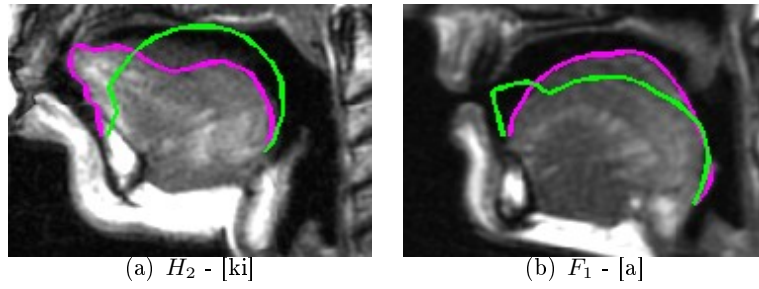


FIGURE 2.24 – Exemples de segmentation erronée : (a) la segmentation est incertaine à cause d'un effet de volume partiel (creux prononcé de la langue dans le plan médio-sagittal) ; (b) l'extrémité gauche de la langue est placée sur l'apex et non plus sur la base du frein lingual (plus bas). La courbe verte donne le contour initial, et la courbe magenta la courbe finale obtenue par notre procédé de segmentation.

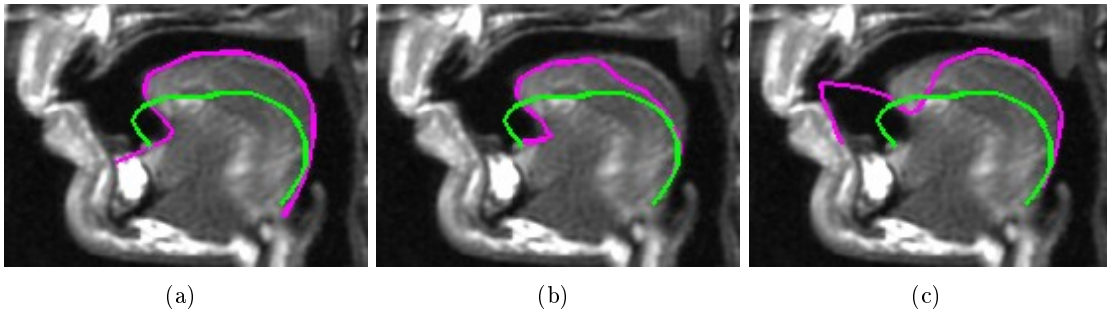


FIGURE 2.25 – Comparaison de notre modèle (a) avec celui de [CV01] (b) et celui de [LKGD07] (c) (locuteur H_1 , son [pu])

- des utilisateurs finaux experts, avec une expérience empirique de l’acquisition des données, mais n’ayant pas toujours bien intégré la notion d’erreur de mesure dans leurs analyses.

Ce cadre nous a permis d’aborder des problématiques classiques, de calibration, de recalage, de segmentation, de modélisation, mais sous une forme exacerbée pour deux aspects. D’une part, la langue, au centre de nos investigations, est l’organe qui présente la mobilité la plus rapide dans le corps humain, et subit des déformations qui sont de l’ordre de sa taille. D’autre part, les modalités en jeu (capteurs EM, US, IRM) sont très faiblement redondantes en utilisation courante.

Ce dernier point nous a amenés à recourir assez fortement à des procédures de calibration, au moyen de dispositifs dédiés, comme le fantôme US/EM que nous avons conçu pour le recalage spatial, ou encore les procédures de synchronisation mises en œuvre. Mais nos contributions ont couvert l’ensemble de la chaîne depuis l’acquisition des données, leur traitement (segmentation), jusqu’à la production de données articulatoires et la confrontation à quelques modèles utilisés en étude de la parole. Une contribution plus transversale a été la détermination d’une incertitude pour l’intégralité du système multimodal, de l’ordre de 2.5 mm. La similarité du système ARTIS avec les systèmes contemporains de nos travaux autorise à leur supposer un même niveau d’incertitude, qui exclut de la négliger dans la phase de modélisation.

Les limitations de ces travaux sont donc en premier lieu ce niveau d’incertitude qui reste élevé, mais est, à notre avis, intrinsèque à l’utilisation de plusieurs modalités, aucune ne délivrant une précision inférieure au demi-millimètre, voire au millimètre. En second lieu, les études sont limitées, avant tout par le temps très conséquent passé entre séances d’acquisition de données et périodes de traitement de ces données. De fait, nous nous sommes essentiellement focalisés sur la langue et les lèvres (via la stéréovision), alors que le larynx et notamment le voile du palais sont aussi mobiles. Par ailleurs, les données articulatoires ont été produites dans le plan médio-sagittal, alors que tout était a priori en place pour aborder une modélisation 3D.

Les progrès récents en acquisition d’images marquent selon nous une rupture avec les principes qui ont fondé le système ARTIS. Trois principales modalités d’imagerie autorisent ainsi aujourd’hui une acquisition d’images dynamiques 3D, autrement appelées images 4D.

La tomodensitométrie X (plus connue sous le terme scanner X, ou CT pour *Computed Tomography* en anglais) permet d’acquérir un volume 3D en une fraction de seconde. Par exemple, le dernier scanner de Toshiba (Aquilion One Vision), permet d’acquérir un volume 3D en 0.275 s. Les techniques actuelles, reprenant la même idée que le fan-beam CT, regroupent les projections mais selon leur synchronisme temporel et proposent une acquisition 4D d’organes subissant un mouvement cyclique (notamment la respiration, comme [LZZ⁺15]). Des études récentes exploitent les techniques d’acquisition comprimée afin de produire un volume par projection dans la séquence d’acquisition CT, mais elles ne sont encore que dans un état très prospectif [CJG⁺14]. Dans tous les cas, le caractère ionisant de l’acquisition scanner semble cependant restreindre cette modalité à des acquisitions de courtes séquences⁹.

L’échographie 4D a bénéficié récemment du développement de transducteurs matriciels et autorise aujourd’hui un flux d’acquisitions 3D à raison de quelques volumes par seconde [BKN⁺16]. Cependant, la sonde étant tenue à la main, les déformations mesurées incorporent le mouvement éventuel de la sonde, qui doit dès lors être traquée, par exemple au moyen de capteurs électromagnétiques [BH11]. Deux autres freins à l’exploitation de l’échographie 4D sont d’une part la déformation que peut entraîner la nécessité d’un bon contact entre la sonde et la peau, et la qualité image qui reste limitée.

Enfin, **l’imagerie par résonance magnétique (IRM)**, bien que théoriquement plus limitée par la physique, a connu ces dernières années une véritable révolution algorithmique qui a permis une réduction spectaculaire des temps d’acquisition. Les techniques les plus récentes associent imagerie parallèle et acquisition comprimée. Une acquisition 3D impose encore un temps de quelques secondes [Hol15], mais la séquence FLASH permet des acquisitions qualifiées de «ci-

9. Par exemple, l’Aquilion One Vision induit une dose de 0.26 mSv par volume reconstruit. Nous rappelons, pour donner un ordre d’idée car la question est plus complexe que cela, que la limite d’exposition naturelle fixée pour la population est de 1 mSv/an

neMRI» 2D à une fréquence de 30 voire 100 Hz [UZV⁺10]. Cette séquence est aujourd’hui la plus prometteuse et est déjà explorée dans le but d’analyser le conduit vocal [LBV⁺16], et même les mouvements de la langue les plus extrêmes comme ceux réalisés par les trompettistes [IFV⁺15].

La perspective à plus court terme paraît donc de continuer à améliorer l’acquisition IRM et mettre au point un protocole permettant d’observer le conduit vocal dans une seule modalité. Cette question ne rentre pas dans nos domaines de compétence. À plus long terme, l’élaboration d’un modèle 3D dynamique du conduit vocal est plus en rapport avec nos intérêts. Il serait intéressant d’adapter au conduit vocal les travaux développés en anatomie computationnelle sur les modèles statistiques non-linéaires de formes [PF15, ZG16]. Cela laisserait espérer en particulier une moins grande sensibilité des descripteurs au placement des points d’amer. Les difficultés seraient de gérer à la fois la grande mobilité intrinsèque de la langue et la grande diversité inter-locuteurs, et de guider la construction du modèle afin que les dimensions réduites puissent être interprétées comme des paramètres articulatoires. Il serait aussi pertinent d’intégrer des données EM, localisant avec une précision croissante ces points d’amer [SBSG17], dans la phase de construction de ces modèles statistiques. Cependant, malgré quelques excursions très exploratoires, l’exploitation de modèles articulatoires 3D, notamment pour la synthèse de sons, est encore éloignée des préoccupations actuelles en étude de la parole. Des progrès restent encore à réaliser au moyen de modèles articulatoires dans le plan médio-sagittal avant d’envisager d’investir une dimension supérieure.

Chapitre 3

Vers une navigation augmentée en neuroradiologie interventionnelle

Sommaire

3.1	Introduction	47
3.1.1	Contexte	47
3.1.2	Contexte médical	48
3.1.3	Contributions	49
3.2	Évaluation clinique de l'imagerie péri-opératoire 3DDSA	51
3.2.1	Premier contexte clinique : traitement endovasculaire des anévrismes intracrâniens	52
3.2.2	Second contexte clinique : traitement des MAV par radiochirurgie	54
3.3	Modélisation et calibration de l'arceau vasculaire	59
3.3.1	Contexte	59
3.3.2	Données capteurs et géométrie idéale de l'arceau	59
3.3.3	Étude et modélisation de l'arceau	63
3.4	Fluoroscopie Augmentée : concept et évaluation	66
3.4.1	Rappel du contexte clinique	67
3.4.2	Principe et prototype de FA	68
3.4.3	Protocole de validation du prototype de FA	68
3.4.4	Résultats et discussion	69
3.5	Reconstruction du micro-guide par stéréovision	71
3.5.1	Contexte	71
3.5.2	Reconstruction stéréo de courbes	71
3.5.3	Principe de l'algorithme proposé	73
3.5.4	Pré-sélection des fragments 3D	74
3.5.5	Ordonnancement des fragments 3D en supports candidats	76
3.5.6	Résultats et discussion	77
3.6	Conclusion et perspectives	79

3.1 Introduction

3.1.1 Contexte

La chirurgie actuelle suit une tendance de plus en plus forte et inéluctable à développer des approches minimalement invasives. Ces nouvelles procédures tirent profit des techniques d'imagerie per-opératoire de plus en plus performantes, mais ont également pour conséquence une perte importante de retour visuel et tactile pour le chirurgien.

Cet environnement crée donc de nouveaux challenges auxquels la Vision Artificielle peut répondre, grâce à la puissance de calcul dont nous disposons aujourd'hui, et malgré le contexte de

temps réel, ou temps au moins clinique [CP10, MIH11]. Cleary et Peters [CP10] identifient quatre technologies-clés qui sous-tendent tout système d’intervention guidée par l’image : la localisation (*tracking*) des outils chirurgicaux et du patient (ou de ses organes), le recalage (*registration*) des différentes données de localisation et d’imagerie dans un repère commun, la visualisation de ces données fusionnées, et le logiciel nécessaire à la conception d’un système de navigation utilisable dans un contexte clinique. Leur point de vue applicatif est cependant réduit à la superposition d’informations 3D sur la vue du chirurgien dispensée par le système d’imagerie interventionnelle. Mirota et al. [MIH11] précisent le contexte de visualisation 2D comme étant à lier à la problématique de la Réalité Augmentée et ajoutent donc comme technologies la reconstruction et le suivi du mouvement des organes et des outils. Ces technologies sont des marches vers ce qui est pour nous l’objectif final à atteindre : travailler non plus dans l’espace de l’imagerie interventionnelle, mais dans un espace entièrement 3D.

Ces points de vue omettent cependant, à notre avis, certaines technologies-clés, qui ne sont qu’évoquées dans ces articles. La première concerne l’acquisition d’imageries pré- ou per-opératoires 3D qui doivent être adaptées à l’intervention [KOJC12]. Certains auteurs pensent même aujourd’hui que cette acquisition doit être repensée et ne plus être générique (comme on utilise un CT ou une IRM aujourd’hui), mais devra cibler l’information à acquérir pendant l’intervention en fonction du contexte [NFF16]. La deuxième concerne la segmentation [LDC⁺13]. Segmenter les organes, les cibles pathologiques, les outils dans les images 3D et interventionnelles 2D, est d’une part souvent une étape nécessaire ou facilitante pour le recalage, mais l’extraction de points d’amer est également une technologie de base pour la reconstruction de surfaces. Par ailleurs, la représentation segmentée d’un organe ou d’une tumeur aide à sa visualisation dans un contexte multimodal, et permet d’en extraire des informations quantitatives (par exemple de taille ou encore de proximité entre l’outil et l’organe). Enfin, la segmentation est un pré-requis à la modélisation, mécanique, biologique, chimique, physiologique. Cette dernière prend une place de plus en plus importante, par exemple pour la prise en compte de l’élasticité des organes, mais également pour mieux contrôler l’interaction physique entre l’outil et l’organe, prévoir le résultat d’une embolisation ou encore prendre en compte le flux sanguin.

Bien que nous ayons abordé certaines de ces problématiques dans un contexte de chirurgie laparoscopique du rein (recalage non-ridige, segmentation et visualisation) [32, 53] et du foie (reconstruction, recalage non-rigide, et modélisation biomécanique) [7, 18, 19, 22], nous n’aborderons dans ce manuscrit que la neuroradiologie interventionnelle qui est notre contexte applicatif de prédilection, mais également car nous avons abordé dans ce contexte la plupart des technologies pré-citées.

3.1.2 Contexte médical

Différentes thèses se sont déroulées dans l’équipe avec pour cadre applicatif la neuroradiologie interventionnelle [Lau96, 71, Anx03, Gor07, Yur14, Del17] (j’ai participé à l’encadrement des quatre dernières). Leurs manuscrits proposent tous à leur façon une introduction à la neuroradiologie interventionnelle. Nous ne reprenons ici que les éléments de contexte essentiels à l’appréhension des travaux repris dans ce chapitre.

La radiologie interventionnelle par cathétérisme est une technique de chirurgie minimalement invasive : une simple incision est faite en général au niveau de l’aîne, par laquelle un *cathéter* (long tube fin et flexible) est inséré dans l’artère fémorale. Le réseau artériel étant totalement connecté, et le sang étant présent dans tout le corps, le médecin, radiologue interventionnel, peut guider le cathéter à travers le réseau artériel pour atteindre un site pathologique. Des traitements peuvent être ainsi appliqués de manière ciblée, en particulier des micro-outils peuvent être amenés et manipulés à distance afin d’opérer.

Le médecin contrôle son geste grâce à une imagerie temps réel qui, dans notre cas est de la radiographie à faible dose de rayons X appelée *fluoroscopie*. La fluoroscopie n’offre cependant qu’une piètre qualité d’image, en deux dimensions (2D). Le médecin dispose également de plusieurs imageries 3D pour analyser la pathologie (voir Figure 3.1). En particulier, l’angiographie rotationnelle 3D (3DRA) peut être acquise pendant l’intervention [10], mais cela requiert

quelques minutes (si on inclut la préparation de l'arceau). Le neuroradiologue interventionnel travaille donc dans un contexte naturellement multimodal où l'imagerie 3D lui permet de saisir, avant, pendant et après l'opération, la forme de la pathologie et l'angioarchitecture liée, mais est statique ; et l'imagerie 2D permet un retour visuel en temps réel sur le geste, la dynamique des micro-outils et du flux sanguin, mais ne peut pas reproduire des informations de structure et de forme essentielles au guidage du geste et à la prise de décision thérapeutique pendant l'intervention.

Les difficultés opératoires se situent donc à deux niveaux. D'une part les structures vasculaires et les pathologies ont souvent une forme très complexe en trois dimensions, dont la fluoroscopie seule ne saurait rendre compte. La Figure 3.2(a) donne un exemple de cas d'anévrisme difficile à analyser en fluoroscopie à cause des superpositions. La 3DRA (Figure 3.2(b)) en revanche montre clairement les relations des artères ainsi que la forme et la position de l'anévrisme. D'autre part, la navigation et le placement des micro-outils (ici un micro-guide dans un micro-cathéter) sont délicats du fait de la grande précision requise et de la complexité du comportement mécanique et dynamique des micro-outils. La fluoroscopie augmentée, contribution de notre travail dont un exemple est montré sur la Figure 3.2(c), facilite cette navigation, grâce à une visualisation en réalité augmentée.

3.1.3 Contributions

La 3DRA a été rendue disponible pour une application en routine clinique en 1997. La recherche d'une intégration et exploitation optimales de cette modalité en neuroradiologie interventionnelle a fondé les travaux que nous reprenons dans ce chapitre.

Un premier ensemble de recherches a eu pour but de bien appréhender son intérêt particulier pour le traitement des pathologies en neuroradiologie interventionnelle. Le cas de l'anévrisme intra-crânien a été traité dans [10, 56] et celui des malformations artério-veineuses lors de la thèse de René Anxionnat [Anx03, 8, 35, 36, 55, 54, 37]. Nous reprenons brièvement les résultats principaux de ces travaux dans la Section 3.2.

Ces travaux mettent en évidence l'importance du recalage géométrique entre le volume 3DRA et les images d'angiographie ou de fluoroscopie 2D. J'avais abordé ce problème dans ma thèse [71] en proposant de calibrer un modèle mécanique de l'arceau vasculaire, vu comme un bras robotisé portant une caméra, afin de relier les données fournies par les moteurs axiaux à la matrice de projection [41]. Ce recalage n'était cependant pas toujours assez précis et nous avons proposé un algorithme de recalage 2D/3D iconique, basé sur une minimisation du flux optique entre la vue projetée de la 3DRA et l'image angiographique 2D [39, 40].

Sébastien Gorges a repris dans sa thèse [Gor07] (contrat Cifre avec la société GE Healthcare) l'approche par calibrage mécanique et l'a affinée en proposant une méthode de calibration par ajustement de faisceaux [33]. La variation des paramètres intrinsèques a ainsi pu être mise en évidence et mesurée, mais au prix d'une procédure lourde. Deux modèles d'arceau, dont un à paramètres intrinsèques constants, plus aisé à calibrer, ont été proposés et validés [33, 51]. Nous reprenons ces travaux dans la Section 3.3. Un fantôme a également été mis au point, associé à une procédure de calibration plus légère, afin de mesurer avec précision la variation des paramètres intrinsèques à travers l'homographie induite dans le plan image [30, 34].

Cette calibration généralisée (capacité de générer une matrice de projection associée à toute configuration de l'arceau vasculaire) a ouvert la voie à des applications multimodales. La première, que nous avons nommée Fluoroscopie Augmentée (FA) consiste à projeter une vue en rendu volumique de la 3DRA sur le flux fluoroscopique utilisé par le médecin pour contrôler son geste. Un exemple en a été donné sur la Figure 3.2(c). Nous en reprenons les éléments de conception essentiels ainsi que l'évaluation pré-clinique que nous avons menée pour la valider [1, 51] dans la Section 3.4, première étape vers ce qui est devenu un produit vendu par la société GE Healthcare (*Innova Fusion*).

Sébastien Gorges avait abordé dans sa thèse des travaux très préliminaires visant à segmenter un micro-guide en temps réel dans les images fluoroscopiques biplan et le reconstruire par stéréovision en 3D, de manière à ce qu'il puisse être visualisé conjointement avec le volume 3DRA. Ces

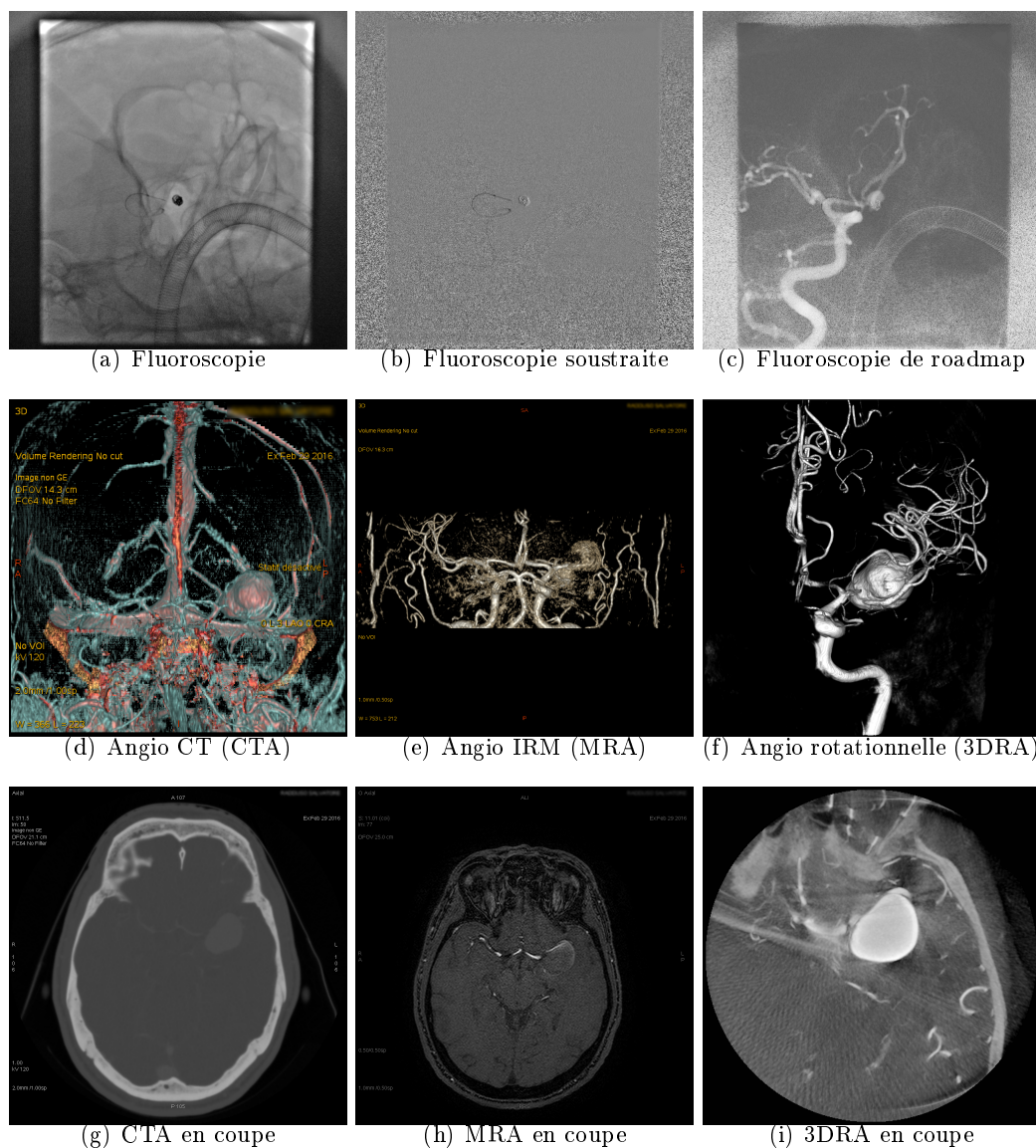


FIGURE 3.1 – Modalités d'imagerie médicale employées en neuroradiologie interventionnelle : en haut, les différents modes de fluoroscopie mettent en évidence les difficultés pour visualiser l'anatomie et les micro-outils ; au centre, les 3 modalités d'angiographie 3D utilisées, chacune présentée avec un rendu volumique : la 3DRA offre la meilleure résolution vasculaire ; en bas, les trois mêmes modalités, en coupe axiale : l'IRM offre la plus riche information anatomique, mais présente des artéfacts de flux (absence de l'anévrisme).

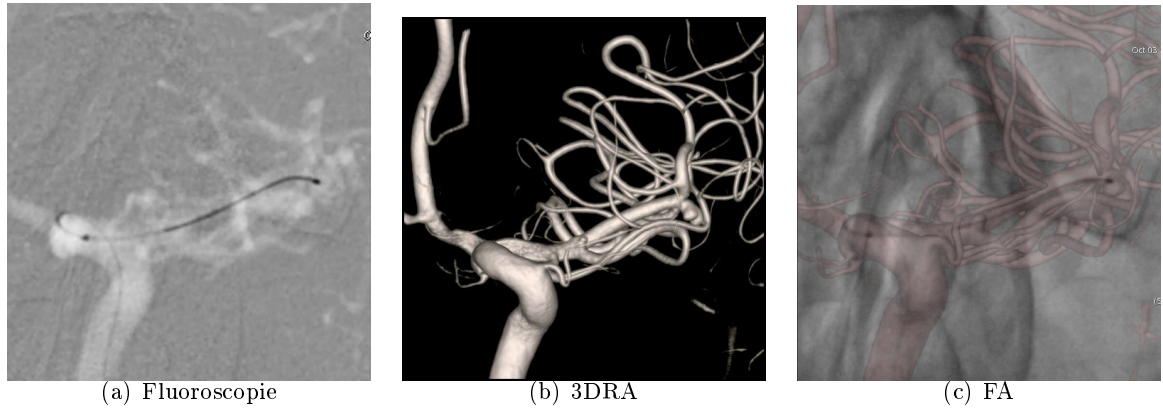


FIGURE 3.2 – Exemple de difficultés opératoires pour emboliser un anévrisme (l’extrémité distale du cathéter est dans la poche anévrismale) : (a) Image de fluoroscopie de roadmap sans visibilité de l’anévrisme perdu dans les superpositions de la branche sylvienne ; (b) le volume 3DRA montre très clairement la forme de l’anévrisme ; (c) l’image en fluoroscopie augmentée (FA), contribution de nos travaux, permet d’intervenir avec une excellente visibilité de la pathologie.

travaux étaient restés embryonnaires et ont été repris en profondeur dans la thèse de Charlotte Delmas [Del17] (contrat Cifre avec GE Healthcare). Nous en synthétisons la partie consacrée à la reconstruction stéréoscopique du micro-guide [15] dans la Section 3.5.

Les travaux revus dans ce chapitre ont été pour l’essentiel effectués dans le cadre de notre collaboration de longue date avec la société GE Healthcare, notamment à travers les contrats Cifre qui ont servi de cadre aux thèses de Sébastien Gorges soutenue en 2007 [Gor07], et celle de Charlotte Delmas, en vue d’être soutenue en 2017 [Del17]. La thèse de René Anxionnat, relative à la Section 3.2 a été soutenue en 2003 [Anx03]. Cette thématique de recherche, y compris mes travaux de thèse, a donné lieu à 10 brevets [57, 58, 59, 60, 61, 62, 63, 64, 65, 66].

Les travaux en segmentation et reconstruction vasculaires feront l’objet du chapitre suivant.

3.2 Évaluation clinique de l’imagerie péri-opératoire 3DDSA

Bien que des systèmes expérimentaux aient été proposés depuis quelques années [SFTP⁺94, TBG⁺96, FFLH97, HCG⁺98] le premier système clinique à proposer une reconstruction 3D suite à une acquisition d’angiographie rotationnelle a été installé en fin d’année 1996 au département de neuroradiologie interventionnelle diagnostique et thérapeutique du CHU de Nancy (système biplan LCN+, GE Healthcare). L’acquisition consistait en une rotation sur 200° à 40°/s du plan frontal, suivi d’un retour à 30°/s, puis d’une deuxième rotation à 40°/s pendant injection intra-artérielle du produit de contraste. La reconstruction reposait ainsi sur 44 images et prenait environ 7 minutes. Du fait des deux rotations, les images étaient soustraites, raison pour laquelle nous l’avons nommée 3D DSA (3D Digital Subtracted Angiography).¹

La 3DDSA était une imagerie nouvelle qui paraissait prometteuse aux ingénieurs. Mais il était nécessaire de bien cerner son intérêt dans la pratique clinique, et notamment sa complémentarité avec les autres modalités d’imagerie employées, notamment l’angiographie numérique soustraite (DSA pour *Digital Subtracted Angiography*), mais aussi le CT et l’IRM, avant d’envisager des outils informatiques et fonctionnalités techniques pertinents pour les médecins. Évaluer une modalité d’imagerie ne fait pas sens dans l’absolu. On ne peut raisonnablement que l’évaluer pour une tâche précise. Nous avons donc réalisé cette étude dans les contextes thérapeutiques du traitement endovasculaire des anévrismes par coils et du traitement des malformations artério-veineuses (MAV) par radiochirurgie.

1. La 3DRA, qui est proposée aujourd’hui par tous les constructeurs, repose sur des acquisitions non soustraites (cf Section 3.3).

3.2.1 Premier contexte clinique : traitement endovasculaire des anévrismes intracrâniens

Contexte

Un anévrisme est une dilatation locale de la paroi vasculaire. Le risque associé est essentiellement hémorragique, conséquence d'une rupture de la paroi ainsi fragilisée. Les anévrismes intracrâniens sont dans une très large proportion sacculaires². Les deux options de traitement sont d'une part la pose par neurochirurgie d'un clip placé à la base de la poche anévrismale pour l'exclure de la circulation sanguine, et d'autre part l'embolisation endovasculaire par déploiement de coils qui viennent progressivement remplir la poche pour empêcher le sang d'y circuler.

La reconstruction 3D des vaisseaux sanguins est particulièrement intéressante dans l'évaluation de cas complexes d'anévrismes intracrâniens. La visualisation précise du collet³ de l'anévrisme et son angioarchitecture⁴ sont des facteurs essentiels pour choisir le traitement (notamment endovasculaire vs neurochirurgical). De plus, sa taille et sa forme guident le choix des coils pour l'embolisation. La DSA par cathétérisme sélectif était, à l'époque de ces travaux, l'imagerie de référence pour évaluer ces facteurs [APC⁺95, KTM⁺96, VRR⁺98, IKY97], mais des tentatives apparaissaient pour exploiter le CT [STH⁺94, ATS95, VSAG95, VRR⁺98] ou l'IRM [IKY97, CJL⁺99]. Cependant, ces modalités ont une résolution insuffisante et une sensibilité plus faible dans la détection des petits anévrismes (diamètre < 3 mm) [STH⁺94, VSAG95, IKY97, CJL⁺99, KTM⁺96, VRR⁺98, YDK99, WTW⁺01].

Matériel et méthode

Notre étude [10] a porté sur 20 patients traités pour hémorragie subarachnoïdienne aiguë consécutive à la rupture d'un anévrisme intracrânien (22 anévrismes, dont 21 traités par voie endovasculaire, et 1 par voie chirurgicale). La comparaison a porté sur trois modalités d'imagerie : la DSA biplan, la séquence d'angiographie rotationnelle, dont certains travaux reconnaissaient l'intérêt [SKO89, HWtG94, TCM⁺96, CWC⁺98], et la 3DDSA, selon deux modalités de visualisation : la vue MIP (*Maximum Intensity Projection*), proche de la vue DSA en ce sens qu'elle n'apporte aucune information de profondeur, et la vue de surface (SSD pour *Surface Shaded Display*) qui la contient, mais au prix d'un seuillage manuel (voir la Figure 3.3).

Deux neuroradiologues interventionnels seniors ont revu tous les cas de manière indépendante dans une première phase, suivie d'une phase de consensus pour les cas de désaccord. L'évaluation a porté sur trois critères : visualisation et position exacte du collet de l'anévrisme ; détermination de la forme de l'anévrisme ; possibilité du traitement endovasculaire. Chacun de ces critères perceptifs a été noté sur une échelle de 3 niveaux : insuffisance d'information, ambiguïté ou indécision, visualisation claire et décision nette. Dans les cas où un traitement endovasculaire a été réalisé (21/22), le logiciel permettait de mesurer les dimensions de la poche anévrismale dans une coupe épaisse reformatée passant par l'axe principal de la poche et contenant le collet (précision des mesures évaluée à 0.5 mm sur un fantôme, voir Figure 3.3(c)) et de déterminer une angulation de travail, en manipulant le volume 3D en rotation, translation et zoom puis de la reproduire sur l'arceau vasculaire (voir Figure 3.3(d)). Par conséquent, deux critères supplémentaires ont été évalués : possibilité de définir et reproduire la vue de travail (oui ou non), et choix du premier coil suite aux mesures de taille réalisées (correct ou incorrect).

Résultats et discussion

Le détail des résultats est disponible dans [10]. La principale conclusion de cette étude a été de mettre en évidence un apport majeur de la 3DDSA dans le traitement des anévrismes

2. Un anévrisme sacculaire présente une excroissance en forme de poche posée sur la surface du vaisseau parent, à la différence d'un anévrisme fusiforme où la dilatation est répartie sur toute la circonférence de l'artère (cas par exemple des anévrismes de l'aorte abdominale).

3. Lieu de connexion de la poche anévrismale sur le(s) vaisseau(x) parent(s)

4. Relations de l'anévrisme avec ses vaisseaux parents

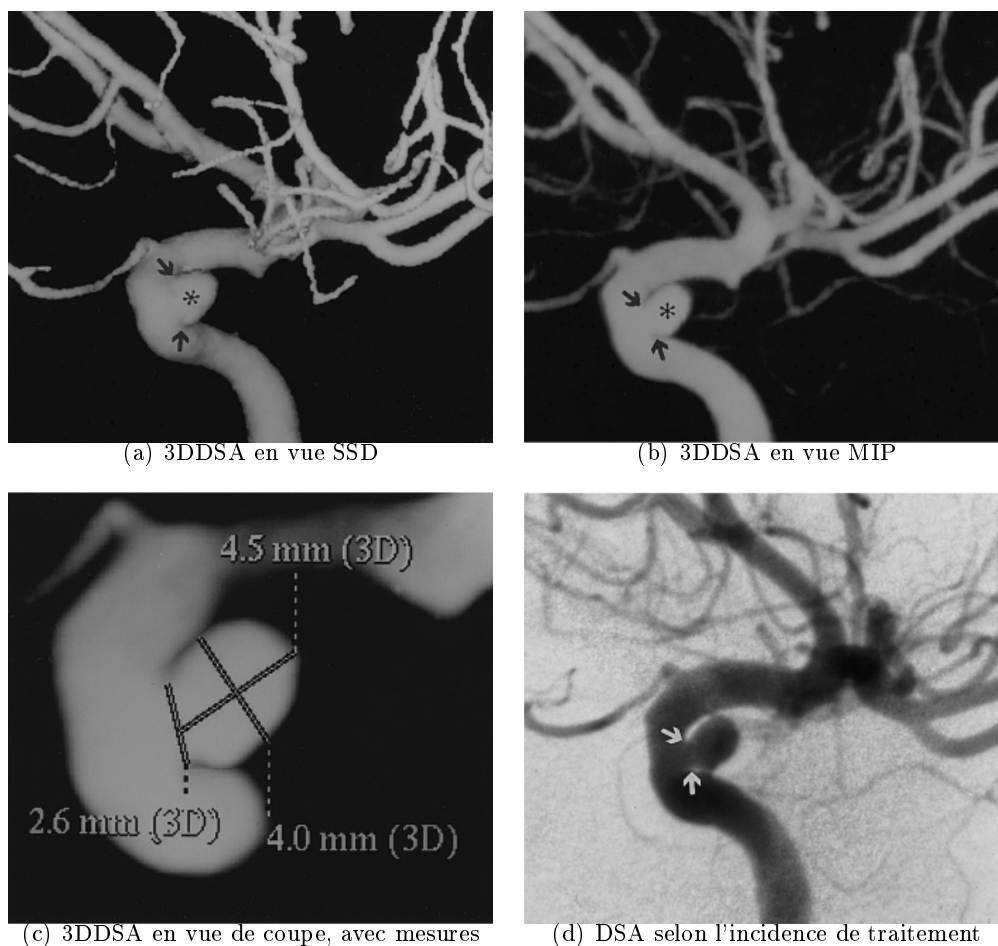


FIGURE 3.3 – Imagerie, visualisation et outils employés pour l'évaluation clinique de la 3DDSA pour le traitement des anévrismes intracrâniens (indiqué par un astérisque en (a) et (b)). (a) 3DDSA en vue SSD ; (b) 3DDSA en vue MIP ; (c) mesure des dimensions de l'anévrisme dans une coupe épaisse reformatée ; (d) DSA. L'angulation de travail a été trouvée en manipulant le volume en vues SSD et MIP ((a) et (b)). Cette angulation a été reproduite pour la vue DSA (d). La vue SSD (a) surestime la taille du collet (flèches), et semble indiquer un cas difficile pour le traitement endovasculaire. Les autres vues montrent en revanche un collet resserré, confirmé par le DSA (d) et le succès du traitement endovasculaire.

intracrâniens, en complément de la DSA. En particulier, la 3DDSA a permis une analyse précise de l'anatomie des anévrismes dans tous les cas, alors que la DSA ne la permettait que dans 4 cas pour l'analyse du collet et 5 cas pour appréhender la forme de la poche. Le mode SSD apporte une vue claire et précise de l'angioarchitecture et de la forme de l'anévrisme. La vue MIP doit en revanche lui être accolée afin de pallier les risques de surestimation de la taille du collet en SSD. Les mesures dans une coupe épaisse reformatée apportent une aide fiable dans le choix du diamètre du premier coil. Ce choix est essentiel car ce coil doit tenir en s'appuyant fermement sur les parois de l'anévrisme, sans le rompre : il forme une grille en 3D qui servira de support aux coils suivants qui viendront remplir la poche de l'anévrisme. Enfin, la possibilité de définir avec confiance une vue de travail en DSA à partir de la 3DDSA est un bénéfice net pour le médecin et le patient, évitant la recherche par tâtonnement qui multiplie les vues obliques, et réduisant ainsi la dose de rayons X reçue et le volume de produit de contraste injecté.

3.2.2 Second contexte clinique : traitement des MAV par radiochirurgie

Contexte

Une malformation artério-veineuse (MAV) intracérébrale est un enchevêtrement localisé d'artères et de veines, formant un *nidus*, qui shuntent le réseau capillaire. L'absence d'irrigation du parenchyme génère de fortes pressions sanguines au niveau de la MAV, menant principalement à un risque hémorragique. En général, le traitement des MAV, notamment complexes, se fait en deux étapes. Une embolisation par voie endovasculaire permet dans un premier temps de réduire la taille du nidus. Le résidu est ensuite oblitéré par radiochirurgie, qui consiste en une seule séance de forte irradiation à travers le crâne intact, en conditions stéréotaxiques⁵.

L'objectif de la radiochirurgie est de délivrer une dose de radiation létale à l'intérieur d'un volume cible, idéalement correspondant au nidus, et, par un fort gradient au niveau de sa surface, une dose très faible à l'extérieur de ce volume. Les causes de l'échec de l'oblitération d'une MAV sont diverses, mais l'irradiation partielle du volume cible en est la principale [GMM⁺98, PFL⁺98, FKL⁺99, CCF⁺03, ZSA⁺03]. La définition précise du volume cible est donc une étape essentielle : il doit inclure la totalité du nidus, et exclure les tissus sains environnants.

Plus précisément, nous avons donc cherché à évaluer la possibilité de définir la cible radiochirurgicale à partir de la 3DDSA. En l'absence de vérité terrain, la référence était le contournage manuel fait par des experts sur des images DSA prises sous diverses incidences. Ce protocole est établi comme un *gold standard* par son utilisation en routine clinique depuis des années avec un très bon taux de succès.

Base de comparaison

L'évaluation d'une segmentation en comparaison à une base d'experts est un problème difficile. Un degré de désaccord peut être défini si on dispose d'une métrique entre courbes [CK01]. Des méthodes statistiques plus sophistiquées ont également été proposées mais elles demandent une assez bonne connaissance a priori sur le processus de segmentation en termes de sensibilité et spécificité [WZWI04]. Dans notre cas, des mesures classiques de comparaison (index DICE, distance moyenne, distance de Hausdorff) ont mis en évidence une forte variabilité dans le détournage manuel des MAV [37]. En l'absence de vérité terrain, notre approche a consisté à modéliser le gold standard par un modèle statistique, issu d'une Analyse en Composantes Principales (ACP), de la variabilité des formes définies par un panel d'experts. L'objet d'étude naturel serait la forme de la MAV en 3D, reconstruite par intersection de cônes à partir des détournages en DSA. Il est cependant reconnu que cette forme est imparfaite, voire grossière, étant notamment incapable de capturer toutes les parties concaves du contour de la MAV [Lau95] (voir Figure 3.4). Nous avons donc pris comme référence les courbes dessinées en 2D par les experts. Pour chaque vue DSA, une ACP est construite à partir des détournages des experts (t vecteurs propres conservés

5. La stéréotaxie, capacité de localiser dans un espace connu l'anatomie du patient, était assurée par un cadre fixé au crâne du patient et portant des marqueurs de localisation. Ces conditions sont aujourd'hui assurées par des procédés de recalage d'images [MTLP12, MBB⁺17].

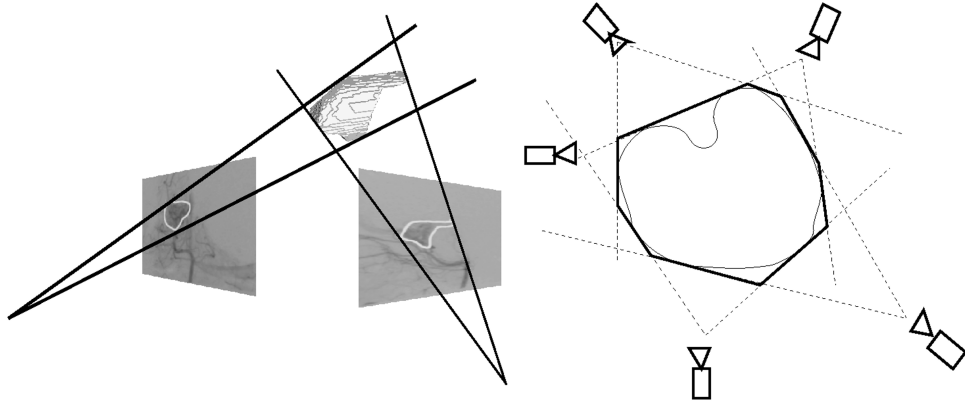


FIGURE 3.4 – Reconstruction de la forme du nidus d’une MAV à partir de contours dessinés sur des images d’angiographie. (gauche) Configuration classique avec deux vues orthogonales. (droite) La forme reconstruite est largement sur-estimée et de nombreuses parties, notamment concaves, ne peuvent être capturées dans la forme, même en multipliant les incidences.

pour 95% de variance expliquée, voir aussi la Section 2.7). La compatibilité d’une courbe test \mathcal{C} avec le modèle ACP est mesurée par le critère [CTCG95] :

$$F_m(\mathcal{C}) = \sum_{i=1}^t \frac{b_i^2}{\lambda_i} \quad (3.1)$$

où b_i est la coordonnée de \mathcal{C} selon le i^e vecteur propre, et λ_i la variance observée selon cet axe propre. Plus exactement, ce critère mesure la distance de Mahalanobis à la moyenne, de la projection de \mathcal{C} sur le sous-espace vectoriel propre associé à l’ACP. Le résidu orthogonal de cette projection est donc implicitement supposé petit. Afin de le vérifier, le résidu attendu r_e est évalué sur la base des M contours par une méthode de Jackknife :

$$r_e^2 = \frac{1}{M} \sum_{i=1}^M r_i^2(\mathcal{C}_i) \quad (3.2)$$

où $r_i(\mathcal{C}_i)$ est le résidu observé sur la courbe \mathcal{C}_i en la projetant sur la base propre d’une ACP réduite, construite en considérant toutes les courbes sauf \mathcal{C}_i . De manière similaire, le résidu de la courbe test \mathcal{C} est mesuré par :

$$r_o^2(\mathcal{C}) = \frac{1}{M} \sum_{i=1}^M r_i^2(\mathcal{C}) \quad (3.3)$$

De manière commensurable avec F_m , la compatibilité entre le résidu attendu et le résidu observé est indiquée par :

$$F_r(\mathcal{C}) = \frac{r_o^2(\mathcal{C})}{r_e^2} \quad (3.4)$$

Le critère final $F(\mathcal{C}) = F_m(\mathcal{C}) + F_r(\mathcal{C})$ suit une loi du χ^2 dont nous avons expérimentalement vérifié qu’elle avait t degrés de liberté (voir [Anx03] pour plus de détails). Le test de conformité d’une courbe \mathcal{C} que nous avons proposé opère donc par un test du χ^2 à t degrés de liberté et un niveau de risque de 5%, qui permet de dire si une courbe donnée est conforme aux détourages des experts dans la modalité de référence. Une modalité de test est évaluée par le pourcentage de conformité des contours que les experts y déterminent.

Nous avons évalué le niveau de conformité naturel de la DSA en calculant par validation croisée quel pourcentage de courbes détourées en DSA était conforme aux $M - 1$ courbes restantes [36]. Le pourcentage assez faible (77.5%) nous a amenés à revoir la définition de notre base d’apprentissage. Il est apparu que les experts exploitaient intuitivement la contrainte épipolaire

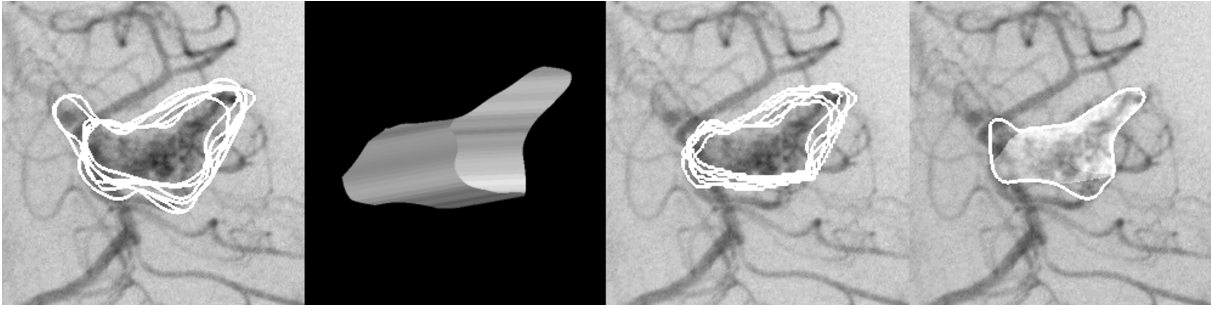


FIGURE 3.5 – (de gauche à droite) 1) Ensemble des contours experts initiaux. 2) Exemple de forme reconstruite en 3D pour le nidus par intersection de cônes pour un détournement expert. 3) Nouvelle base de contours issus de la reprojexion des formes reconstruites dans cette vue. 4) Exemple de contour initial (blanc) superposé avec la reprojexion de la forme reconstruite (en transparence). La contrainte épipolaire exclut certaines parties du contour initial. Cet effet est assez répétable chez les experts, ce qui conduit à une nouvelle base présentant moins de variabilité.

en détournant largement une portion incertaine du contour dans une vue, et comptant sur une vue auxiliaire où la définition est plus claire, pour corriger ce contour (voir Figure 3.5). Nous avons donc employé comme nouvelle base les contours reprojétés de la reconstruction visuelle de la cible issue des contours initiaux. La variabilité apparaît visuellement réduite, ce qui est confirmé par un pourcentage de conformité interne à cette nouvelle base de 83%.

Dispositif expérimental

Notre étude [8] a porté sur 11 patients, disposant chacun d’une 3DDSA ainsi que de 2 à 5 vues DSA (frontale, latérale et d’éventuelles obliques) calibrées grâce au cadre de stéréotaxie, pour un total de 37 vues DSA sur la base de données. La MAV a été détournée dans chaque vue DSA par un panel de 3 experts (deux venant du CHU de Nancy, et le dernier du Karolinska Institute de Stockholm, tous trois neuroradiologues interventionnels seniors). Chaque expert a réalisé deux détournements espacés dans le temps, avec présentation aléatoire des patients. Ainsi la base de référence comportait 6 contours par vue DSA.

La construction de l’ACP requiert que toutes les courbes soient rééchantillonnées sur le même nombre de points [CTCG95], en correspondance univoque basée sur leurs indices. Un grand soin doit être apporté à cette étape d’appariement, qui doit respecter les raisons, souvent complexes, physiques, biologiques, ou intellectuelles à l’origine des schémas de variabilité observés [DTC⁺02]. Nous avons analysé que dans le cas des MAV, un contour alterne entre zones de consensus entre experts, lorsque le contour est clair, visuellement et/ou d’un point de vue médical, et zones d’incertitude, notamment quand le contour est obstrué par la superposition avec une veine de drainage ou par du matériel d’embolisation. En l’absence d’un critère numérique clair à exploiter [Tag99, CK01], nous avons réalisé une mise en correspondance semi-interactive grâce à une interface permettant d’indiquer par des segments les interfaces entre zones de consensus et zones d’incertitude (voir Figure 3.6). La correspondance est ensuite basée sur l’abscisse curviligne sur chaque zone ainsi déterminée.

Afin d’éviter un biais potentiel en lien avec le protocole de détournement, chaque 3DDSA était présentée aux experts sous forme de vue MIP dans la même incidence que les vues DSA acquises pour le patient. Les 3 experts ont à nouveau détourné, avec le même outil, les MAV dans ces vues. Afin de compenser l’absence d’information de flux, disponible en DSA, une visualisation de la 3DDSA en rendu volumique était proposée en complément afin de permettre d’analyser l’angioarchitecture en 3D.

Par ailleurs, l’IRM était une autre modalité d’imagerie envisagée pour remplacer la DSA. Nous avons également évalué l’IRM à travers notre protocole. Le protocole T1 injecté employé ne permet pas de générer des vues MIP de bonne qualité. La cible a donc ici été dessinée par un



FIGURE 3.6 – Interface logicielle d'appariement semi-interactif de courbes. Après que tous les contours ont été superposés à la vue DSA, les segments en noir sont placés manuellement afin d'indiquer les frontières entre les zones de consensus entre experts et les zones de forte variabilité (souvent pour divergence d'avis médical). Les segments gris matérialisent, sur un exemple entre deux courbes, l'appariement effectué en se basant sur l'abscisse curviligne au sein de chaque ensemble de segments de courbe homologues.

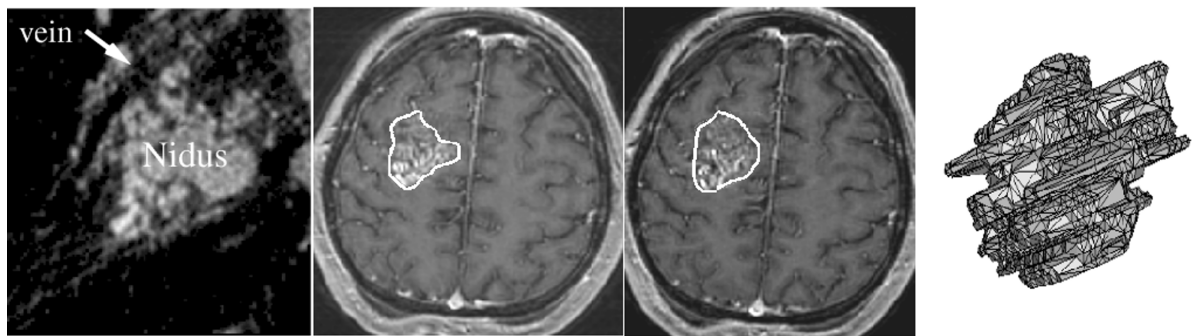


FIGURE 3.7 – Détourage d'une MAV en IRM (pondération T1, injectée), coupe par coupe en coupes axiales. (gauche) Coupe axiale à travers la MAV dans la 3DDSA correspondante. (deux du milieu) Contours de la MAV dans deux coupes IRM successives. (droite) Contour reconstruit en 3D. Les difficultés de détourage apparaissent dans l'aspect chaotique de la surface reconstruite, qui fait apparaître nécessaire un outil de régularisation de la forme dans le sens longitudinal.

détourage coupe par coupe en coupes axiales (voir Figure 3.7).

Résultats et discussion

Les détails des résultats sont disponibles dans [8, 36]. En moyenne, le taux de conformité de la 3DDSA est bon (69%), mais un peu en-dessous du taux de référence (84%, interne aux images DSA). Il est toutefois nettement supérieur à celui de l'IRM (44%). Il apparaît cependant que ces MAV étaient toutes complexes, voire extrêmement difficiles. En se restreignant aux MAV raisonnablement complexes, les taux de conformité sont en revanche très proches pour la 3DDSA (82% pour une référence de 87%), mais restent bas pour l'IRM (60%).

Ces expériences nous ont menés à réfuter l'IRM comme présentant trop de risques d'erreur pour le détourage des MAV. En revanche, la 3DDSA apparaît plus fiable et nous avons préconisé de l'employer en complément de la DSA. Nous avons ainsi développé, lors du contrat d'ingénieur de Frédéric Speisser, un logiciel graphique dédié au détourage multimodalité des MAV [54] permettant de mieux appréhender l'impact d'une vue DSA sur la forme finale, et donc assurer une meilleure cohérence des contours entre les vues (voir Figure 3.8). Nous avons également présenté

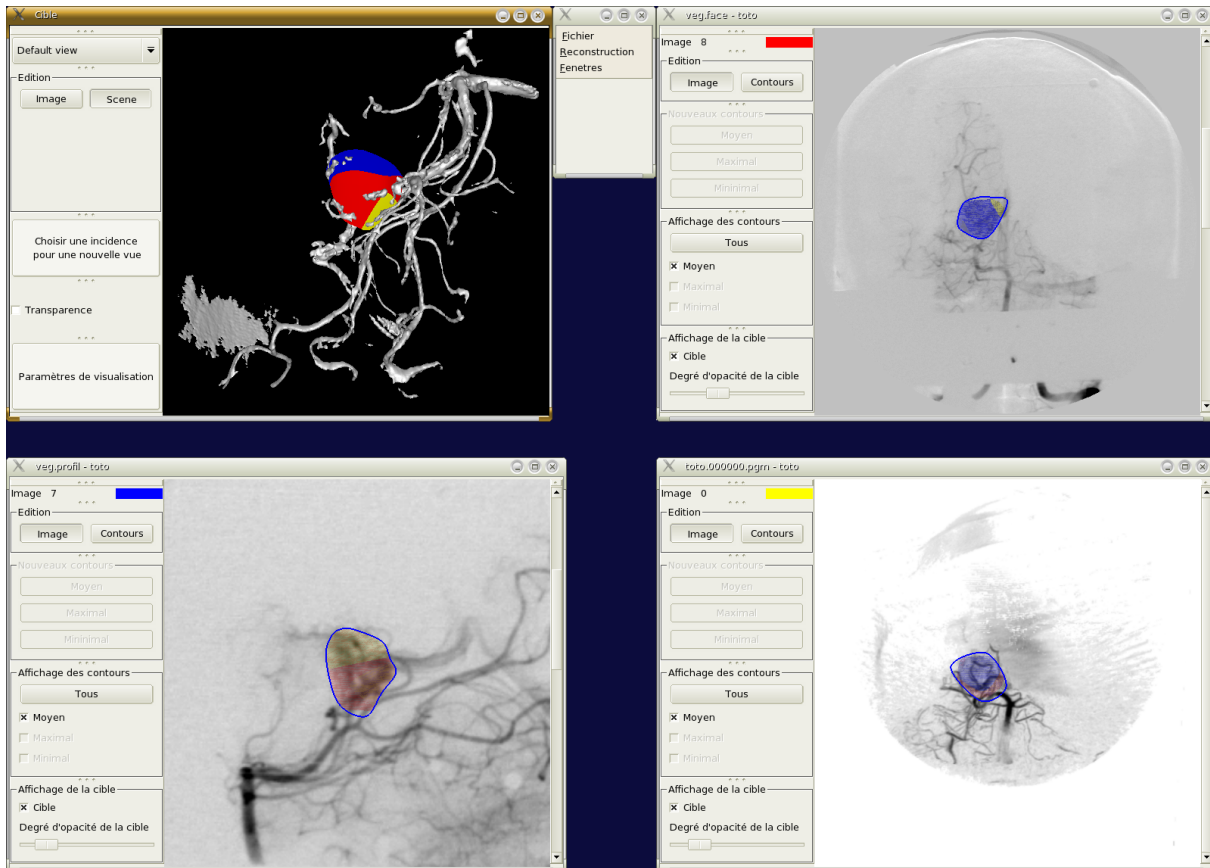


FIGURE 3.8 – Le logiciel InterStax permet de détourner le nidus d’une MAV à partir de vues DSA (haut, droite et bas, gauche) et 3DDSA en MIP (vue "virtuelle", bas, droite). La contribution de chaque contour 2D apparaît en couleur sur la forme finale qui peut être reprojetée en 2D (en transparence) ou superposée à la 3DDSA en vue de surface (haut, gauche).

des résultats préliminaires dans [35] d'un affinage de la forme du nidus, en 3D, par le biais d'une modélisation par contour actif maximisant l'intégrale du gradient de la 3DDSA sur la surface de la forme [LM99b]. Cette procédure a l'avantage de diminuer le volume du nidus, et donc la dose radiochirurgicale reçue par le patient, et, puisqu'elle est automatique, diminue la variabilité (améliore l'objectivité ?) du détournage. Nous avons pu montrer que les contours obtenus ainsi étaient indiscernables des contours experts, au sens de notre test de conformité.

3.3 Modélisation et calibration de l'arceau vasculaire

3.3.1 Contexte

Nos travaux précédents ont éclairé la complémentarité des acquisitions 3DDSA avec les images 2D DSA en neuroradiologie interventionnelle. Une vue plus complète et informative de la pathologie peut donc être proposée au médecin en affichant de manière conjointe les images sous ces deux modalités [51, RHM⁺11]. Satisfaire ce besoin clinique passe par la résolution d'un problème de recalage vasculaire 2D/3D. Des revues récentes témoignent de la diversité des techniques proposées depuis que cette question a suscité l'intérêt des chercheurs il y a 20 ans [MTLP13, MBB⁺17]. Nous avons d'ailleurs été les premiers à proposer une approche par recalage iconique [40]. Cet intérêt s'explique aussi parce que les vaisseaux sanguins sont présents dans tous les organes, souvent en profondeur, proches des pathologies qu'ils alimentent, et présentent une forme complexe qu'il est aisé d'identifier dans plusieurs modalités par injection. Ce sont donc d'excellents amers pour un recalage multi-modalités [MBB⁺17] (voir également [MTLP12] pour un point de vue élargi au recalage 2D/3D).

Toutes ces méthodes reposant sur le contenu des images, déterminent le recalage par optimisation d'une fonction de coût. De nombreuses procédures d'optimisation ont été testées, mais la plupart nécessitent une initialisation, et toutes bénéficient de la connaissance d'un recalage initial assez précis pour faire face aux nombreux optima locaux générés par la structure arborescente des vaisseaux (i.e. la qualité de l'alignement s'améliore localement dès que deux vaisseaux, même non homologues, se superposent en projection). Cette question de l'initialisation est rarement évoquée, et est effectivement difficile dans le cas général, ou plutôt trop dépendante de l'application. Dans le cas d'une procédure utilisant un arceau vasculaire (ce qui est le cas de la neuroradiologie interventionnelle), la chaîne image (couple tube rayons X-détecteur image plan) est portée par un bras motorisé indiquant par des capteurs axiaux les angles de rotation de l'arceau. De plus, la table est elle aussi motorisée en translation. L'exploitation de ces données capteurs a d'abord été faite dans le contexte de la cardiologie interventionnelle pour reconstruire par stéréoscopie l'axe 3D du vaisseau à traiter par stent et déterminer ainsi une vue optimale pour la quantification du degré de sténose sans distorsion projective [DRG94]. Différentes variantes de ce modèle isocentrique originel peuvent être obtenues en relâchant successivement les différentes hypothèses qui le composent, la variante la plus performante étant celle qui les remet toutes en cause [CNMR02]. Dans notre thèse, nous avons suivi une approche purement analytique via des méthodes de métrologie par la vision appliquées à des acquisitions d'une mire de calibration. Nous avons ainsi démontré l'écart entre le comportement mécanique de l'arceau et sa géométrie idéale. Nous avons alors proposé un nouveau modèle liant les paramètres de la projection aux données capteurs [41].

Ce modèle souffrait néanmoins encore de certains a priori, notamment l'hypothèse que les paramètres intrinsèques étaient indépendants de l'orientation de l'arceau, alors que nous avons mis en évidence une petite déformation mécanique de l'arceau sous son propre poids. Par ailleurs, l'arrivée des détecteurs plans, plus légers et dépourvus de distorsion optique, ont justifié de reprendre cette étude sur le nouvel arceau vasculaire en 2004.

3.3.2 Données capteurs et géométrie idéale de l'arceau

Tout notre travail de modélisation s'est concentré sur l'arceau frontal pour deux raisons : d'une part la grande majorité des salles d'angiographie sont équipées d'une machine monoplan,

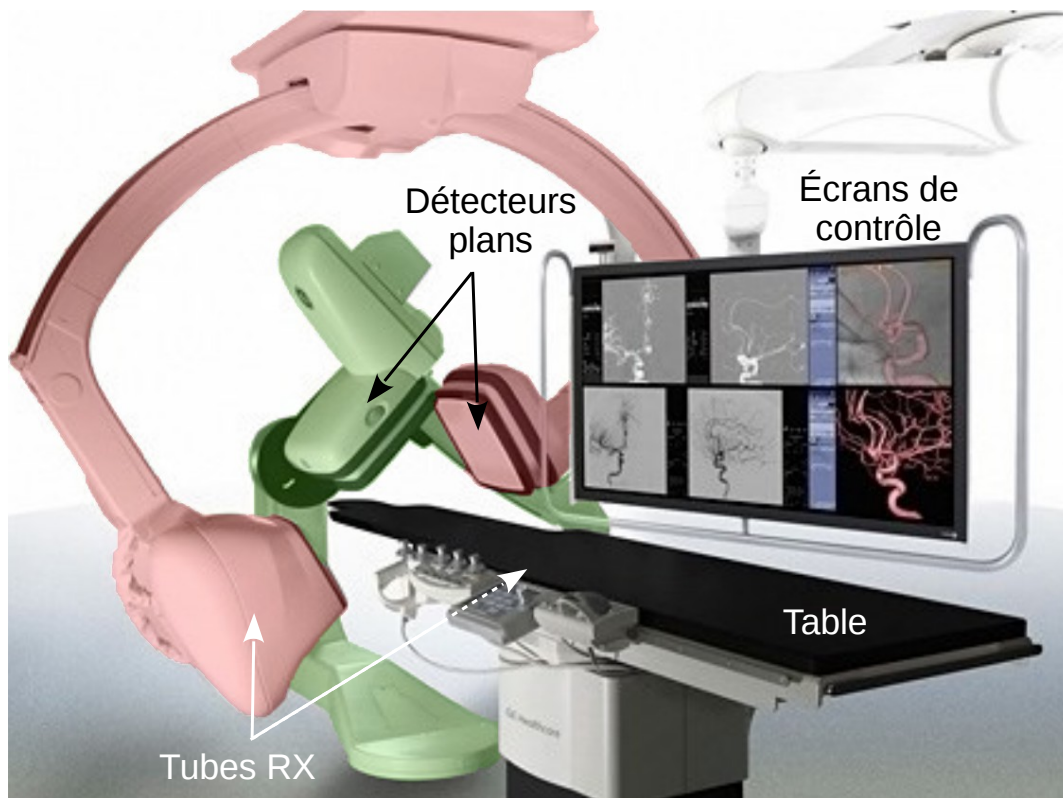


FIGURE 3.9 – Exemple de machine d'angiographie biplan. L'arceau latéral est surimprimé en rouge et l'arceau frontal en vert (le tube rayons X (RX) est caché par la table d'opération).

et d'autre part le plan latéral, lorsqu'il existe, suit la même géométrie que le plan frontal, mais est connu pour être mécaniquement moins précis par design. Une machine biplan est présentée sur la Figure 3.9.

Les données capteurs

Les capteurs peuvent être séparés entre capteurs de position et capteurs d'état.

La position exprime en premier lieu l'incidence de la vue obtenue en modifiant l'orientation de l'arceau (voir Figure 3.10). Cette incidence est repérée par deux angles anatomiques : LAO/RAO (*Left/Right Antero-posterior Orientation*) et CRA/CAU (*Cranio/Caudal angle*). Ces angulations sont reproduites mécaniquement grâce à deux moteurs axiaux dont les angles de rotation sont reportés par les valeurs de capteurs P (LAO/RAO) et C (CRA/CAU). Le modèle géométrique idéal est ainsi composé de deux rotations ; la première selon un angle P autour d'un axe longitudinal (allant du bas vers le haut du patient, autrement connu comme la direction axiale en imagerie 3D), et la deuxième selon un angle C autour d'un axe latéral (allant de la gauche vers la droite du patient, autrement connu comme la direction sagittale en imagerie 3D). Idéalement, ces deux axes sont orthogonaux et se coupent en un point appelé *isocentre*, censé se projeter au centre de l'image quelle que soit l'incidence. Cependant, dans cette configuration, l'arceau limite l'amplitude des mouvements longitudinaux de la table, ce qui est gênant pour les interventions hors du crâne (rachis pour ce qui est la neuroradiologie interventionnelle, mais également en cardiologie, en abdominal, ou pour des interventions sur les vaisseaux périphériques). Un troisième moteur est donc présent qui permet la rotation de l'arceau sur un axe vertical (de l'arrière vers l'avant du patient, autrement connu comme la direction coronale en imagerie 3D), dont l'angulation est repérée par le capteur L. Cet axe est supposé orthogonal aux deux autres, passant par l'isocentre. Une valeur de L non nulle a pour conséquence de rendre caduque l'équivalence ($P=LAO/RAO$) et ($C=CRA/CAU$) et se traduit par une rotation dans le plan image. L'amplitude de cette rotation se déduit facilement du triplet de valeurs (P,C,L), sous l'hypothèse

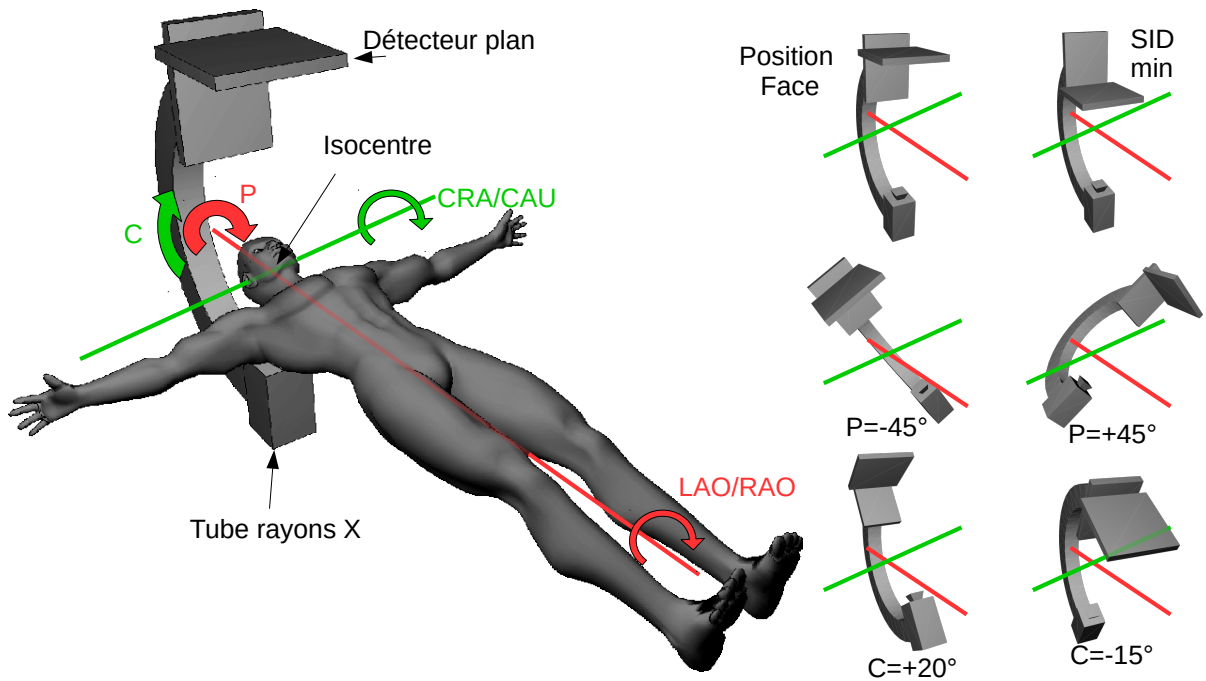


FIGURE 3.10 – Une incidence définie par les angles anatomiques LAO/RAO et CRA/CAU est atteinte en faisant varier les angles mécaniques P et C de l’arceau. De plus, le détecteur plan peut être rapproché ou éloigné du tube rayons X, faisant ainsi varier la SID.

de validité de la géométrie idéale, ce qui permet de redresser automatiquement l’image.

Par ailleurs, le détecteur plan est posé sur un rail motorisé qui permet d’en modifier la distance au tube rayons X. Cette distance est repérée par un capteur appelé SID (*Source to Image Distance*). L’intérêt est de rapprocher le capteur image le plus possible de la tête du patient afin de réduire le flou de foyer (phénomène de pénombre). La distance entre l’isocentre et le foyer du tube rayons X est connue par la valeur (fixe) SOD (*Source to Object Distance*).

La position $P=C=0$ est appelée position antéro-postérieure, ou position de face. Lors d’une intervention, l’angle LAO/RAO (P si $L=0$) varie entre -100° et 100° et l’angle CRA/CAU (C si $L=0$) varie entre -25° et 35° . La SID varie entre 95 et 118 cm. Les capteurs angulaires offrent une précision de 0.1° et la SID 2 mm.

Si nous mettons de côté les différents paramètres (kV, mA, etc...) qui déterminent la qualité de l’image, le seul facteur ayant un impact sur la géométrie de l’image est le FOV (*Field Of View*). Cette valeur était difficile à maîtriser avec les amplificateurs de brillance, notamment à cause de la présence de distorsions. La situation s’est considérablement simplifiée avec l’arrivée des détecteurs plans : une image occupe soit l’intégralité de la matrice de 1000×1000 pixels, soit une zone centrale de 500×500 pixels. Par ailleurs la taille pixel est parfaitement connue (taille d’un micro-capteur sur la matrice du capteur image, multipliée par le zoom numérique) : nous la notons s_p , chaque pixel étant carré par conception. La taille pixel pour un FOV usuel de 20 cm est donc de 0.2 mm.

Le modèle sténopé et sa calibration

Nous supposons le lecteur familier avec le modèle sténopé classiquement utilisé en vision artificielle. Les références à [Fau93, HM95, HZ00] seront bénéfiques pour une éventuelle remise en mémoire, tout comme les thèses réalisées dans l’équipe sur ce sujet [Lau96, 71, Gor07, Del17] (les deux dernières traitant plus directement d’un système à capteur image plan). Nous nous concentrons ici sur son lien avec les données capteurs et le modèle idéal de l’arceau.

La formation d’image rayons X obéit en très bonne approximation à une projection centrale [Web12]. Un point 3D $Q = (X, Y, Z)$ de l’objet se projette au point 2D $q = (u, v)$ dans le

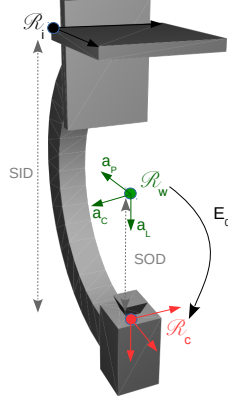


FIGURE 3.11 – Le modèle de Dumay, basé sur la géométrie idéale de l’arceau [DRG94]. Le repère mon \mathcal{R}_w (en vert) est le repère centré sur l’isocentre, et ayant pour axes les axes de rotation autour du P, du C et du L (respectivement a_P, a_C, a_L). Le repère caméra \mathcal{R}_c (en rouge) est centré sur le point focal (point d’émission des rayons X) et des axes parallèles à ceux du repère image \mathcal{R}_i (en noir) pour les deux premiers, le troisième formant une base orthogonale avec eux.

plan image via la matrice de projection \mathbf{P} de taille 4×3 telle que :

$$\underline{q} = \mathbf{P}\underline{Q} = \mathbf{K}\mathbf{E}\underline{Q} \quad (3.5)$$

où le souligné indique le passage des coordonnées euclidiennes aux coordonnées homogènes. La matrice \mathbf{K} des paramètres intrinsèques prend la forme générique

$$\mathbf{K} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.6)$$

où α_u et α_v donnent la distance focale en pixels (valeurs égales si les pixels sont carrés) et (u_0, v_0) est le point principal. La matrice \mathbf{E} des extrinsèques exprime le changement de repère entre le repère monde et le repère caméra et se décompose donc en une rotation \mathbf{R} et une translation T :

$$\mathbf{E} = [\mathbf{R}|T] \quad (3.7)$$

Le modèle génératif de Dumay [DRG94] s’appuie aussi sur ce modèle projectif et considère donc l’arceau vasculaire comme une caméra portée par un bras robotisé. Du côté des paramètres intrinsèques nous avons ainsi $\alpha_u = \alpha_v = \text{SID}/s_p$ et le point principal est placé au centre de l’image. Le repère monde est le repère dit *isocentrique* dont l’origine est l’isocentre, supposé être un point fixe pour toute rotation, et les axes sont les trois axes anatomiques identifiés dans le modèle idéal décrit ci-dessus aux trois axes de rotation de l’arceau (voir Figure 3.11). En position de face, la matrice extrinsèque est ainsi :

$$\mathbf{E}_* = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\text{SOD} \end{bmatrix} \quad (3.8)$$

Et les extrinsèques pour une incidence quelconque (P,C,L) s’en déduisent par combinaison avec les trois rotations axiales correspondantes (L, puis P, puis C dans l’ordre) :

$$\mathbf{E} = \mathbf{E}_* \begin{bmatrix} \cos(C) & 0 & -\sin(C) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(C) & 0 & \cos(C) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(P) & -\sin(P) & 0 \\ 0 & \sin(P) & \cos(P) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(L) & -\sin(L) & 0 & 0 \\ \sin(L) & \cos(L) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.9)$$

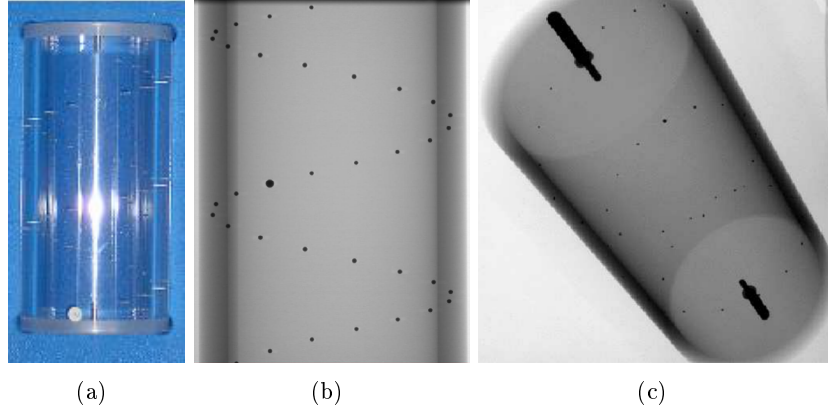


FIGURE 3.12 – Mire de calibration utilisée pour les images rayons X, composée de billes de plomb disposées en hélice sur un cylindre en plexiglas (a). En position de face (b), les billes peuvent être indexées suivant leur position verticale dans l'image. La plus grosse bille indique l'indice médian. En position oblique (c), le modèle de Dumay génère une projection approximative qui permet l'appariement des billes.

La matrice de projection peut également s'estimer si on connaît un ensemble d'appariements entre N points 3D $\{Q_i\}$ et N points 2D $\{q_i\}$ [Fau93]. Ces conditions sont remplies avec une mire de calibration telle que celle conçue pour calibrer les positions où des images sont prises lors d'une rotation de l'arceau en vue d'une reconstruction 3DRA [RPTP93] (voir Figure 3.12). Cette mire est composée de billes en plomb disposées en hélice sur un cylindre de plexiglas. En position de face, les billes peuvent ainsi être indexées, en 3D suivant leur position le long de l'axe du cylindre, et en 2D suivant leur position verticale dans l'image (voir Figure 3.12(b)). Une bille plus grosse sert de référence pour les indices. Dans sa thèse, Sébastien Gorges a étendu l'utilisation de cette mire pour toute position, même oblique, en initialisant l'appariement grâce à la projection approximative fournie par le modèle de Dumay. Quelle que soit la position, cette mire permet une calibration de la matrice de projection avec une erreur de reprojection de 0.1 ± 0.01 pixel. Un calcul direct permet d'estimer la matrice de projection à partir des appariements 2D/3D, mais il a été établi que faire suivre cette initialisation d'une minimisation itérative de l'erreur quadratique moyenne de reprojection \mathcal{E}_r en améliorerait le résultat [Fau93, RPTP93] :

$$\mathcal{E}_r(\mathbf{P}) = \frac{1}{N} \sum_i |\underline{\mathbf{P}Q_i} - q_i|^2 \quad (3.10)$$

où le second souligné indique le passage des coordonnées homogènes aux coordonnées euclidiennes.

Il est possible de décomposer une matrice de projection en paramètres intrinsèques \mathbf{K} et extrinsèques \mathbf{E} [Tos87]. Ainsi, à partir d'une image de la mire de calibration, et connaissant la matrice extrinsèque \mathbf{E}_* pour la position de face, la matrice $\mathbf{D} = \mathbf{E}_*^{-1}\mathbf{E}$ caractérise le déplacement de l'arceau dans l'incidence courante (voir Équation 3.9). Le principe d'étude des mouvements de l'arceau consiste donc à étudier cette matrice de déplacement pour des mouvements axiaux de l'arceau, c'est-à-dire des mouvements selon un seul axe de rotation, les autres angulations étant mises à zéro.

3.3.3 Étude et modélisation de l'arceau

Calibration multi-images

L'indépendance théorique des paramètres intrinsèques et extrinsèques n'est plus valable en présence de bruit dans les données. Il est ainsi établi depuis longtemps que des variations importantes dans ces paramètres peuvent s'observer pour des faibles variations de l'erreur de reprojection autour de son minimum [PS90, PH02]. Dans ce cas, les variations des paramètres

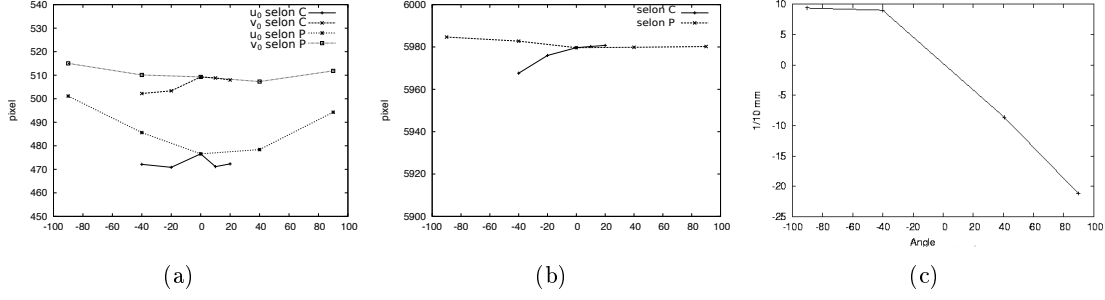


FIGURE 3.13 – Effets de déformations mécaniques de l'arceau en fonction de l'angle (en degrés). (a) Variations du point principal (u_0, v_0) , en pixels. (b) Variations de la focale α , en pixels. (c) Variations de l'amplitude du vissage λ , en dixièmes de mm, selon P.

intrinsèques sont compensées par des erreurs sur les paramètres extrinsèques, dont l'interprétation mécanique est alors délicate. La fiabilité des mesures sur les intrinsèques, et donc sur les extrinsèques, est cependant améliorée lorsqu'on considère plusieurs images de calibration pour lesquelles les intrinsèques peuvent être considérés constants [LVD98, PH02]. Nous ferons référence à cette technique par l'expression calibration multi-images.

Nous inspirant de [LVD98] nous avons proposé dans [33, 34] de procéder à une calibration multi-images de la mire. L'arceau étant en position fixe, M images de la mire sont acquises, en déplaçant celle-ci en rotation et translation entre deux acquisitions. L'erreur multi-images suivante est alors minimisée par l'algorithme de Levenberg-Marquardt :

$$\mathcal{E}_m(\mathbf{K}, \{\mathbf{E}_i\}_{1 \leq i \leq M}) = \frac{1}{M} \sum_i \mathcal{E}_r(\mathbf{K}\mathbf{E}_i) \quad (3.11)$$

Cette méthode s'apparente à l'ajustement de faisceaux [TMHF99], avec l'hypothèse simplificatrice que les points 3D sont connus. La minimisation porte donc sur $3 + 6M$ paramètres : 3 paramètres intrinsèques communs (α, u_0 et v_0) et 6 paramètres extrinsèques pour chaque position de la mire. Nous avons déterminé expérimentalement qu'au-delà de $M = 30$ images les valeurs des paramètres intrinsèques se stabilisaient, avec un impact en reprojection inférieur à 0.5 pixels. De plus, avec cette valeur de $M = 30$, nous avons estimé par bootstrap sur les données réelles et par simulation de Monte-Carlo que les limitations en pratique sur les déplacements de la mire (il faut voir assez de billes) n'avaient pas d'impact sur la qualité de la calibration et que l'écart-type était de 1.8, 1.3 et 1.5 pixels pour respectivement u_0, v_0 et α , soit environ 10 fois moins que par la méthode de calibration mono-image (voir [Gor07] pour plus de détails).

Analyse axiale

Cette procédure a été appliquée pour calibrer deux ensembles d'incidences, regroupant chacun des échantillons d'un mouvement axial : le premier \mathcal{P}_P pour la rotation autour du P ($P \in \{-90^\circ, -40^\circ, 0^\circ, 40^\circ, 90^\circ\}$, et $C=0$) et le deuxième \mathcal{P}_C pour la rotation autour du C ($C \in \{-40^\circ, -20^\circ, 0^\circ, 10^\circ, 20^\circ\}$ et $P=0$). Dans les deux cas, la SID était fixe à 118 cm. Une calibration multi-images a été réalisée pour chaque incidence.

Les variations observées sur les paramètres intrinsèques restent dans les limites de la précision de la mesure, ce qui autorise à les considérer constants, sauf pour u_0 lors de la rotation autour de l'axe P (voir Figure 3.13(a)). Dans ce cas, nous avons évalué qu'un polynôme du troisième degré en P était le plus adapté pour modéliser sa variation [33, Gor07] :

$$u_0(P) = \sum_{i=0}^3 \mu_i P^i \quad (3.12)$$

Pour une incidence donnée dans \mathcal{P}_P ou \mathcal{P}_C le déplacement \mathbf{D} de l'arceau peut être déterminé avec précision en combinant les extrinsèques \mathbf{E} en cette incidence avec ceux de l'incidence de face \mathbf{E}_* (voir Équation 3.9)

$$\mathbf{D} = \mathbf{E}_*^{-1} \mathbf{E} \quad (3.13)$$

Cette manipulation ne fait cependant sens que si le repère monde est le même pour les extrinsèques. Nous avons ainsi pris soin d'inclure dans chaque ensemble d'images pour chaque incidence de \mathcal{P}_P et \mathcal{P}_C une image où la position de la mire reste la même : nous utilisons les extrinsèques estimés pour cette image lors de la calibration multi-images (voir Équation 3.11).

Chaque déplacement \mathbf{D} est idéalement une rotation soit autour de l'axe du P (le même pour chaque incidence dans \mathcal{P}_P , soit autour de l'axe du C (le même pour chaque incidence dans \mathcal{P}_C). Nous déterminons la rotation la plus proche en exprimant \mathbf{D} sous forme de vissage [33] :

$$\mathbf{D} = [\mathbf{R}_\theta | (\mathbf{I} - \mathbf{R}_\theta) \mathbf{O} + \Delta T] \quad (3.14)$$

où \mathbf{R}_θ est une rotation d'angle θ autour de l'axe v , \mathbf{O} est un point sur l'axe de rotation (nous prenons le plus proche de l'origine) et $\Delta T = \lambda v$ est une translation résiduelle, parallèle à l'axe de rotation. Idéalement, nous avons pour \mathcal{P}_P (respectivement \mathcal{P}_C) $\theta = P$ (resp. C), v est l'axe longitudinal (resp. sagittal), \mathbf{O} est l'isocentre et $\lambda = 0$. Nous avons ainsi pu vérifier que les capteurs étaient fiables ($0.26^\circ[0.18^\circ]$ d'erreur moyenne, l'écart-type est entre crochets), les axes de rotation étaient stables ($0.06^\circ[0.02^\circ]$ de variation pour P, $0.48^\circ[0.30^\circ]$ pour C), et que le vissage était nul pour C ($|\lambda| < 0.2$ mm). En revanche, les centres et axes de rotation doivent être calibrés pour les deux axes et le phénomène de vissage est important pour la rotation autour du P (voir Figure 3.13(c)). Nous avons à nouveau expérimentalement déterminé qu'un polynôme du troisième degré modélisait bien la dépendance de ce paramètre avec P (voir [33, Gor07] pour plus de détails) :

$$\lambda = \sum_{i=0}^3 \lambda_i P^i \quad (3.15)$$

Deux modèles mécaniques d'arceau

Au vu de cette analyse, nous avons proposé dans [33] deux modèles d'arceau, avec pour limites à ce travail que l'axe de rotation du L n'était pas inclus et que nous avons fait l'hypothèse que les rotations autour du P et du C étaient indépendantes. Sous cette hypothèse, les extrinsèques pour une incidence quelconque connue par les angles P et C sont donnés en modifiant l'Équation 3.9 par :

$$\mathbf{E}(P, C) = \mathbf{E}_* \mathbf{D}_r(C) \mathbf{D}_v(P) \quad (3.16)$$

où $\mathbf{D}_v(P)$ est un vissage composé d'une rotation d'angle P (donnée capteur) autour d'un axe dirigé selon v_P et passant par O_P , et d'une translation parallèle à v_P d'amplitude $\lambda(P)$ donnée par l'Équation 3.15, et $\mathbf{D}_r(C)$ est une rotation d'angle C (donnée capteur) autour d'un axe dirigé selon v_C et passant par O_C . Les intrinsèques à SID constante sont α , v_0 constants et u_0 fourni par l'Équation 3.12. On obtient ainsi le modèle à 26 paramètres :

$$\mathcal{M}_m(\phi_m) \quad \text{avec} \quad \phi_m = \{\mathbf{E}_*; v_C, O_C; v_P, O_P, (\lambda_i)_{i=0..3}; (\mu_i)_{i=0..3}, v_0, \alpha\} \quad (3.17)$$

Calibrer ce modèle nécessite une acquisition similaire à celle utilisée pour l'analyse axiale (besoin équivalent à une calibration multi-images par position), en prenant soin d'inclure pour chaque incidence une image de la mire dans une position commune à toutes les incidences. Cette procédure est lourde et nous avons donc également inclus dans nos investigations un modèle légèrement simplifié, à intrinsèques constants (u_0 est constant) à 23 paramètres :

$$\mathcal{M}_f(\phi_f) \quad \text{avec} \quad \phi_f = \{\mathbf{E}_*; v_C, O_C; v_P, O_P, (\lambda_i)_{i=0..3}; u_0, v_0, \alpha\} \quad (3.18)$$

Dans ce cas, la calibration ne requiert par incidence, qu'une image de la mire dans une position fixe et commune. Le modèle \mathcal{M}_f est moins précis mécaniquement que \mathcal{M}_m , mais, outre sa calibration extrêmement plus abordable dans un contexte clinique, nous pouvons nous attendre à ce que l'hypothèse de constance du point principal n'ait que peu d'influence sur la qualité de la reprojction.

Calibration des modèles

Le modèle \mathcal{M}_f correspond à une calibration multi-images identique à [LVD98]. Nous employons donc la même méthode qui résout par la méthode de Levenberg-Marquardt le problème de minimisation suivant :

$$\phi_f = \operatorname{argmin} \mathcal{R}_f \quad \text{avec} \quad \mathcal{R}_f = \frac{1}{N} \sum_{(P,C) \in \mathcal{P}_P \cup \mathcal{P}_C} \mathcal{E}_r(\mathbf{E}(P, C)) \quad (3.19)$$

où les ensembles \mathcal{P}_P et \mathcal{P}_C sont les mêmes que précédemment et leur union contient N incidences.

Il n'y a pas de difficulté à adapter cette procédure à la calibration de \mathcal{M}_m (le nombre d'images par incidence est le même), mais il est apparu que la faible influence de la variation de u_0 sur la fonction de coût en rendait l'estimation imprécise. Nous avons donc procédé par relaxation de la calibration du modèle \mathcal{M}_f en résolvant le problème suivant :

$$\phi_m = \operatorname{argmin} \left\{ \mathcal{R}_f + \frac{\gamma}{N} \sum_{(P,C) \in \mathcal{P}_P \cup \mathcal{P}_C} \mathcal{R}_m(P, C) \right\} \quad (3.20)$$

Où \mathcal{R}_f est similaire à l'expression dans l'Équation 3.19, en prenant en compte la variation de u_0 et où $\mathcal{R}_m(P, C)$ est le résidu d'une calibration multi-images dans l'incidence (P,C), dans laquelle les différents extrinsèques \mathbf{E}_i sont les seules variables, le reste étant fixé par la valeur courante de ϕ_m . Nous avons fixé γ à 1000.

Afin d'alléger la calibration d'un modèle à intrinsèques variables, nous avons décrit dans [30] une mire planaire fixée au détecteur plan. La variation des intrinsèques implique une variation de l'image de cette mire. Nous avons montré que l'image subissait une homographie dont les paramètres sont en bijection avec les variations des paramètres intrinsèques. Nous avons pu retrouver les mêmes effets qu'observés sur les intrinsèques avec une calibration multi-images. Nous n'insérons cependant pas de description plus détaillée de ce dispositif car celui-ci n'est resté qu'au niveau de prototype et aurait demandé d'être construit avec plus de précision mécanique pour en tirer tous les bénéfices. Nous référons le lecteur à la publication citée ci-dessus ou encore à la thèse de Sébastien Gorges [Gor07] pour de plus amples détails.

Évaluation des modèles

La précision des deux modèles a été mesurée en comparant la matrice de projection prédite avec celle mesurée au moyen de la mire pour diverses incidences, y compris des incidences hybrides (i.e. avec P et C non nuls) [33, 51]. Sans surprise, le critère de reprojection utilisé n'a pas fait apparaître de différence notable, et les deux modèles présentent une erreur de reprojection de 0.46 pixels à SID fixe, 84% des incidences étant prédites avec une précision inférieure. La variation de SID entraîne une augmentation de l'erreur de reprojection jusqu'à 1.2 mm (nous avons mesuré la répétabilité de l'arceau à 0.6 mm d'erreur de reprojection). Il est probable que le phénomène de vissage, observé pour une rotation autour du P et lié au poids de l'arceau, varie en fonction de la SID. L'impact clinique de ce niveau d'erreur a été mesuré par un test visuel basé sur des images de fantôme vasculaire comme non gênant. La Figure 3.14 montre la superposition d'un volume 3DRA sur une image d'angiographie du fantôme vasculaire avec ce niveau d'erreur (voir [Gor07] pour plus de détails).

3.4 Fluoroscopie Augmentée : concept et évaluation

La capacité de réaliser un recalage 2D/3D fiable, rapide et précis ouvre la voie à des applications cliniques nombreuses. Dans cette section, nous présentons la Fluoroscopie Augmentée (FA), qui implante le concept de Réalité Augmentée adapté à la radiologie interventionnelle. Nous verrons dans la section suivante une autre application, qui propose une reconstruction 3D du micro-guide à partir d'une acquisition biplan calibrée.

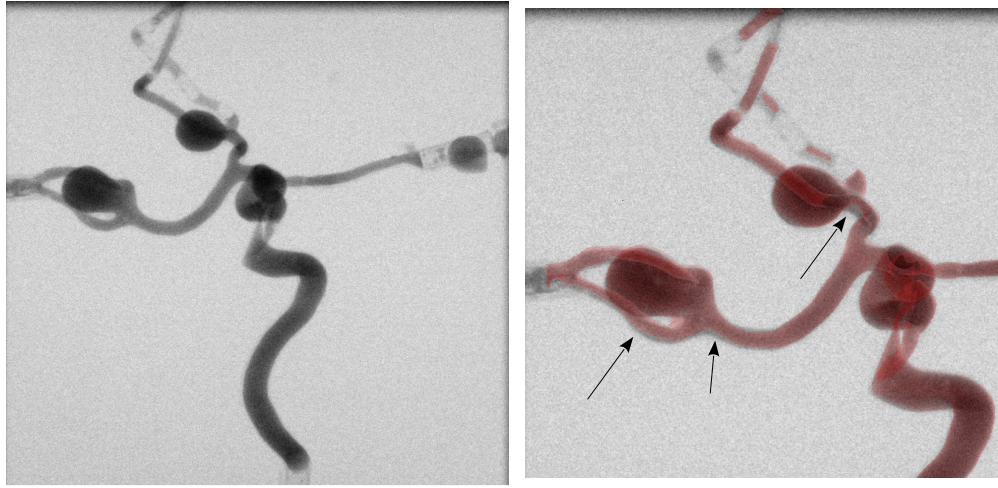


FIGURE 3.14 – Image d’angiographie de fantôme vasculaire (à gauche) sur laquelle a été superposé un volume 3DRA (en rouge, sur l’image zoomée de droite). L’incidence est donnée par les angles $P=-40.4^\circ$ et $C=28.8^\circ$. L’erreur de reprojection mesurée est de 1.2 mm (erreur maximale observée), et est responsable de petits décalages estimés non gênants par les neuroradiologues interventionnels.

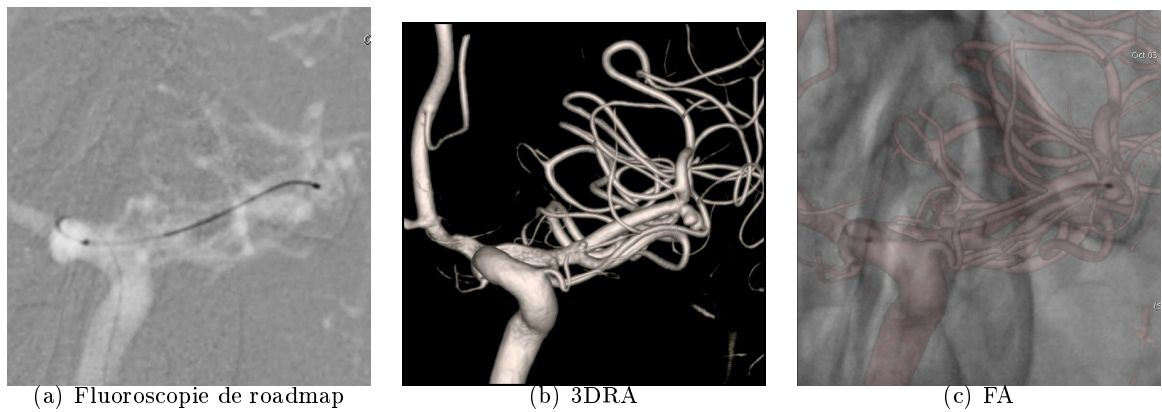


FIGURE 3.15 – Apport visuel de la fluoroscopie augmentée (FA). (a) Image de fluoroscopie de roadmap traditionnellement utilisée pour la navigation des micro-outils en neuroradiologie interventionnelle. (b) Image de 3DRA recalée par notre recalage basé machine. (c) Image de FA obtenue en fusionnant la vue 3DRA recalée avec l’image de fluoroscopie non soustraite.

3.4.1 Rappel du contexte clinique

Dans la phase de navigation, le neuroradiologue interventionnel utilise des images fluoroscopiques pour avoir un retour visuel sur son geste. Ces images sont cependant d’un type particulier appelé *roadmap 2D* : l’image présentée est une image sans injection, soustraite d’une image masque acquise avec opacification des artères. Sans soustraction, les structures osseuses gêneraient la visualisation du micro-outil, certes opaque, mais très fin. Et sans opacification, les vaisseaux resteraient invisibles alors que tout le travail de la navigation consiste à s’y repérer. Une image de roadmap est présentée sur la Figure 3.15(a) (nous y reprenons la Figure 3.2). Malgré la bonne visualisation des artères, le manque de relief rend les superpositions illisibles, par exemple au niveau de l’anévrisme qui est très clairement visible sur la 3DRA. Par ailleurs, lors d’une navigation complexe, il peut être nécessaire de changer d’incidence de travail, ce qui nécessite de procéder à une nouvelle injection de produit de contraste pour former l’image de roadmap.

3.4.2 Principe et prototype de FA

La FA suit le principe de réalité augmentée (RA) appliquée aux images per-opératoires, ces dernières étant augmentées d'informations péri-opératoires, notamment les images 3D acquises avant ou en début d'intervention, éventuellement travaillées, par exemple pour en extraire des structures d'intérêt.

Le recalage est à la base de ce type de systèmes, et bien souvent les travaux de validation se focalisent sur la précision du recalage. En établir une mesure qui conserve sa pertinence en conditions cliniques est difficile. Une étape intermédiaire, mais toujours longue et complexe, consiste alors en des expériences sur fantôme [PFJ05, CMM⁺16], cadavre [FDM⁺04, PBC⁺06, NHT10] ou sur animal [MCK⁺13]. Dans notre cas, nos travaux de la section précédente ont établi cette précision sous l'hypothèse, réaliste, que le patient reste immobile et les vaisseaux sanguins sont rigides.

Un système de réalité augmentée comporte cependant d'autres éléments dont il est aussi important d'évaluer la pertinence et l'adéquation à la pratique clinique des médecins [KOJC12, LDC⁺13]. Il a par exemple été prouvé que le mode de visualisation peut influencer sur la qualité du geste [JEH02, EJJH⁺04, SBHN06]. Certains critères cliniques sont quantitatifs, comme la réduction du temps de procédure, de dose, de volume de produit de contraste [SBHA05]. Mais notre objectif ici est de développer un prototype qui s'intègre sans heurt au protocole et la pratique interventionnels.

La FA a été conçue comme une imagerie de roadmap améliorée dans laquelle le relief des artères apparaît. Un principe similaire (nommé *3D Roadmap* par les auteurs) a été présenté de manière indépendante par [SBHA05]. L'implantation de ce principe dans un prototype clinique a cependant demandé de faire des compromis afin de préserver l'intégrité du système d'acquisition et de traitement des images et garantir la sûreté de l'intervention. Le plus important a été de déléguer toute l'implantation logicielle et matérielle à notre partenaire industriel (GE Healthcare), notre rôle se cantonnant à celui de conseil et aide à la conception.

La solution retenue s'appuie sur un PC branché en parallèle du système d'angiographie, avec un retour écran en salle. Les algorithmes nécessaires au traitement des données fluoroscopiques, ont été implantés au sein d'un logiciel de visualisation sur ce PC. Des raisons techniques ont conduit à considérer la fluoroscopie non-soustraite comme image de base, sur laquelle le volume 3DRA est projeté. Le volume 3DRA est pré-traité par un seuillage manuel afin d'extraire, par *marching cubes*, la surface des vaisseaux sanguins sous forme de maillage triangulaire. OpenGL est utilisé pour le rendu de cette surface qui est fusionné avec l'image fluoroscopique. Un exemple de rendu est reproduit sur la Figure 3.16 : deux niveaux de transparence étaient proposés aux médecins, l'un permettant de mieux percevoir le relief des vaisseaux sanguins (Figure 3.16(a)), et l'autre préservant mieux la visibilité du micro-outil en fluoroscopie (Figure 3.16(b)). Les données capteurs de la table n'étant pas assez précises [41], il a été posé comme contrainte que la table ne devait pas être déplacée entre l'acquisition 3DRA et la session FA afin de pouvoir réaliser un recalage basé sur la modélisation de l'arceau seul. Cette condition n'est pas handicapante, n'ayant pour conséquence qu'un petit excentrage de l'anévrisme dans l'image. Par ailleurs, nous pouvons supposer sans difficulté que le patient ne bouge pas, celui-ci étant placé sous anesthésie générale lors d'un traitement endovasculaire. Le rendu est automatiquement mis à jour si l'arceau est déplacé pendant l'examen. Enfin, nous supposons que les vaisseaux ne se déforment pas sous l'effet de l'insertion des micro-outils, ce qui n'est pas strictement correct, mais cette hypothèse est également faite par la fluoroscopie de roadmap, le masque contenant les vaisseaux étant fixe et formé hors présence des micro-outils.

3.4.3 Protocole de validation du prototype de FA

Notre objectif était d'évaluer d'une part si la FA, conçue proche de la fluoroscopie de roadmap, ne perturbait pas la pratique des médecins, voire leur apportait des informations pertinentes et fiables pendant l'intervention. D'autre part, dans la logique d'un prototype à visée industrielle, il s'agissait aussi de valider les choix de conception du système, et identifier des points d'amé-

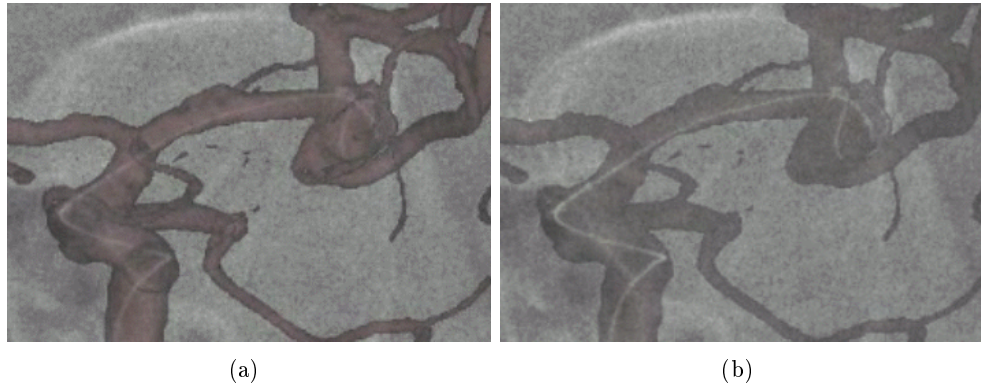


FIGURE 3.16 – Les deux modes de rendu de notre prototype de FA. (a) Superposition d’une vue de surface en transparence : le relief est bien perçu mais peut gêner la visibilité de l’outil (ligne blanche). (b) Un deuxième mode propose une transparence plus forte, ce qui améliore la visibilité de l’outil au détriment de la perception du relief.

lioration pour favoriser l’intégration du système dans la pratique médicale. Nous avons donc développé une grille d’évaluation perceptive afin de recueillir le ressenti des neuroradiologues interventionnels. Cette grille d’évaluation permet de caractériser la qualité du recalage dans les conditions cliniques (respect des hypothèses), la perception des outils dans les images et l’apport potentiel de la technique par rapport à la pratique courante, à travers cinq items :

1. la qualité du recalage : notée mauvaise, acceptable, ou bonne ;
2. l’erreur de recalage perçue : < 3 mm, < 2 mm, < 1 mm, < 0.5 mm ;
3. la position de l’outil en 2D : très souvent, souvent, parfois ou jamais hors des vaisseaux sanguins. Le neuroradiologue doit indiquer la cause, selon lui, d’un décalage éventuel ;
4. la visibilité de l’outil dans l’image : mauvaise, acceptable ou bonne ;
5. la position de l’outil en 3D (capacité de comprendre où l’outil est positionné en 3D) : mauvaise, acceptable ou bonne.

Cette grille a été remplie par deux neuroradiologues interventionnels seniors pour trois patients. Trois modes de visualisation ont été ainsi annotés : la fluoroscopie de roadmap, et les deux modes de transparence proposés pour la FA (Figure 3.16).

3.4.4 Résultats et discussion

Le détail des résultats est donné dans [51] et une analyse plus approfondie est disponible dans [Gor07]. Nous ne reprenons ici que les points essentiels à retenir de cette étude. La qualité du recalage a été jugée inférieure au millimètre (voire mieux). Le cathéter a été perçu une fois légèrement en-dehors du vaisseau mais c’était également le cas sur la fluoroscopie de roadmap. La FA n’a jamais été notée inférieure à la fluoroscopie de roadmap. Sans surprise, la FA en forte transparence permet une meilleure visibilité de l’outil alors qu’une transparence moins prononcée aide à mieux comprendre l’angioarchitecture aux bifurcations. En termes de conception de produit, il est apparu que la visualisation 3D devait être améliorée et que le logiciel devait offrir la possibilité de moduler la transparence dans la vue fusionnée. Ces remarques ont été prises en compte par les ingénieurs de GE Healthcare et un produit, nommé *Innova Fusion* a été lancé en 2007, et est utilisé en routine clinique depuis 2011. La Figure 3.17 compare la visualisation de notre prototype clinique avec celle d’Innova Fusion. Notre étude ne repose que sur un nombre très faible de patients, mais le principe d’évaluation a été repris, et bien sûr amendé, par la société GE Healthcare lors de son processus de validation clinique impliquant bien plus de patients. Des fonctionnalités ont été ajoutées, comme la possibilité d’afficher une pathologie en vue avant comme arrière afin d’en mieux analyser l’angioarchitecture (voir Figure 3.18). Un retour de notre expérience de ce développement de la recherche à l’utilisation en routine clinique a été publié dans une note éditoriale [1].

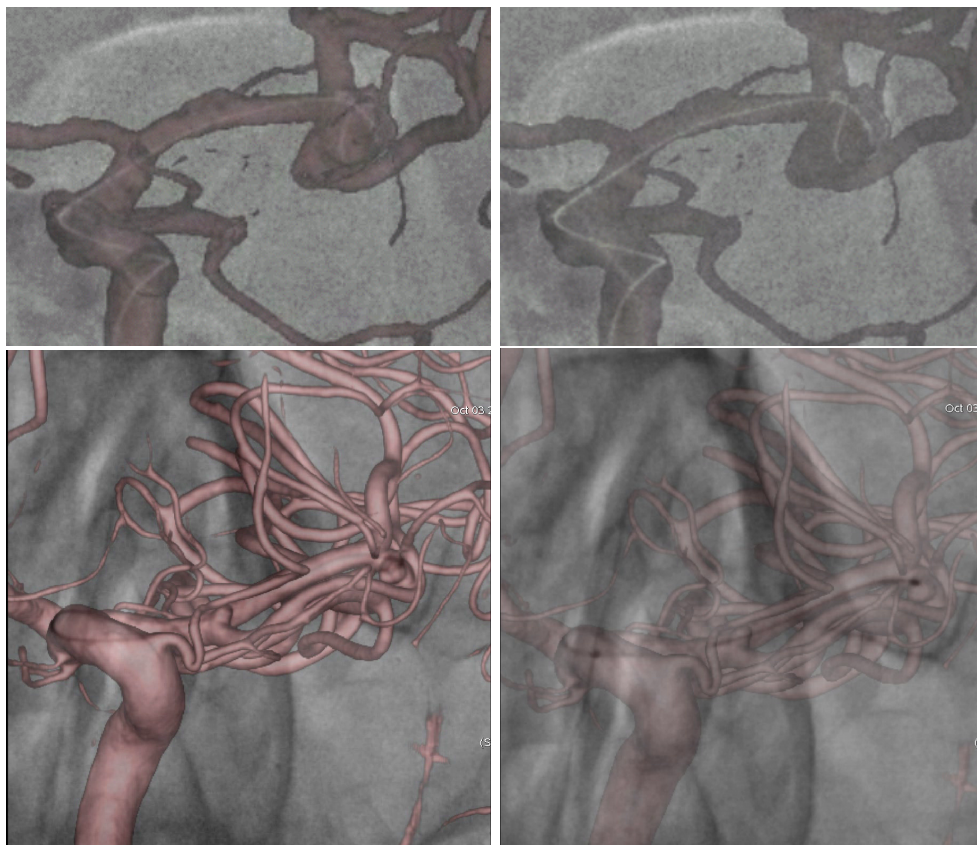


FIGURE 3.17 – Comparaison de notre prototype (ligne du haut) avec le produit Innova Fusion (ligne du bas). Les patients sont différents. La colonne de gauche montre les vues au relief prononcé, et la colonne de droite les vues en transparence forte. La qualité des vues offertes par le produit est nettement meilleure.

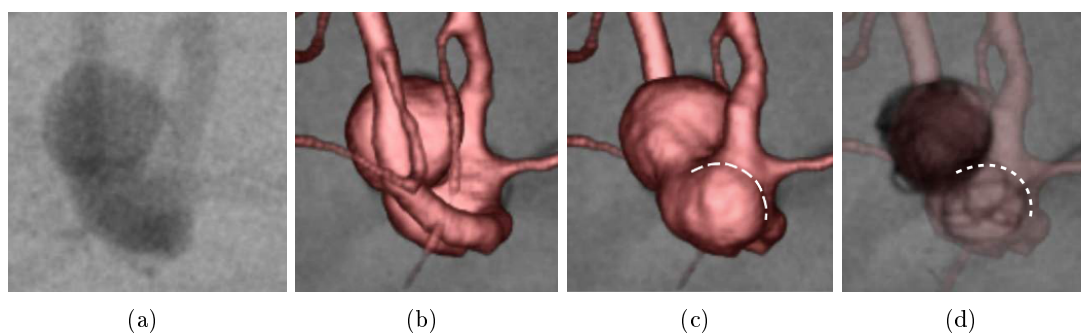


FIGURE 3.18 – Intérêt de la fonctionnalité du produit Innova Fusion autorisant des vues avant et arrière de la pathologie. (a) Vue en fluoroscopie injectée d'un anévrisme multilobé situé sur une bifurcation complexe ; (b) vue avant de l'anévrisme : les superpositions des vaisseaux afférents en gênent l'analyse ; (c) vue arrière de l'anévrisme : la limite d'embolisation est plus facile à déterminer afin de préserver le flux sanguin dans les artères afférentes ; (d) vue FA en transparence où le neuroradiologue interventionnel peut vérifier le bon positionnement de ses coils. Nous avons ajouté les pointillés blancs à la main afin de faciliter la comparaison des images.

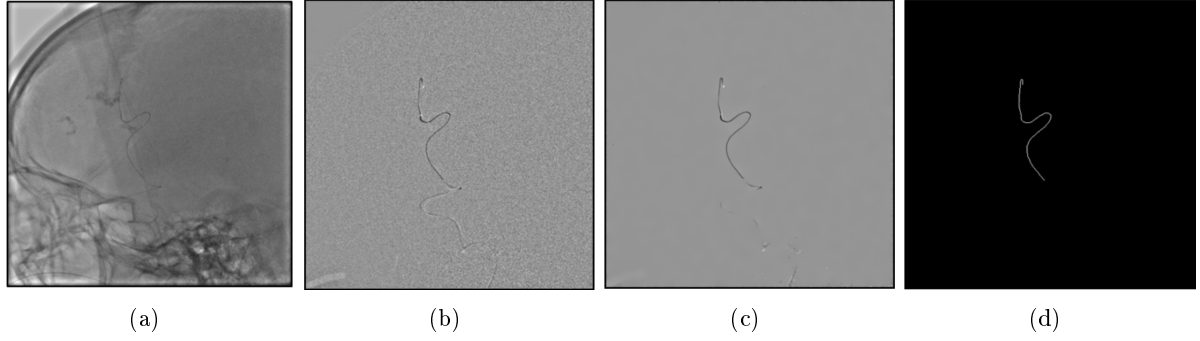


FIGURE 3.19 – Illustration des étapes conduisant à la segmentation du micro-guide : (a) image initiale ; (b) image transformée : image soustraite avec stabilisation de la variance afin d’obtenir un bruit de fond uniforme ; (c) image débruitée : filtre par *Coherence-enhancing Diffusion* [Wei99] modifié afin d’appliquer un fort filtrage isotrope dans les zones uniformes (voir [Del17, chapitre 2] pour les détails) ; (d) image segmentée : seuillage de l’image filtrée suivi d’une squelettisation.

3.5 Reconstruction du micro-guide par stéréovision

3.5.1 Contexte

Sébastien Gorges avait abordé la question de la détection du micro-guide dans sa thèse [Gor07, chapitre 6]. L’objectif premier était d’en améliorer automatiquement la visibilité dans un affichage dans une vue FA avec relief, mais rapidement l’idée de le reconstruire en 3D a été explorée [Gor07, chapitre 7]. La détection proposée se faisait par suivi de courbe dans les images successives de la séquence fluoroscopique en mode biplan. La courbe dans les images frontale et latérale était régularisée par une spline 2D, ce qui nous plaçait dans les conditions supposées par [BvdKvW⁺03]. Nous avons donc évalué les performances de cette méthode sous l’hypothèse que les deux arceaux frontal et latéral étaient calibrés par notre recalage basé machine. Une précision évaluée à $1.3 \text{ mm} \pm 0.3 \text{ mm}$ par simulations de Monte Carlo et des premiers résultats cliniques ont été jugés encourageants.

Plusieurs groupes, dont GE Healthcare, ont montré un grand intérêt pour les méthodes de détection, sous forme de segmentation ou simplement d’amélioration de la visibilité du cathéter et du guide dans les images fluoroscopiques. Le lecteur intéressé peut se référer à [BVG09] pour les travaux antérieurs à 2009 et à [Del17, chapitre 2] pour une revue des travaux plus récents. En début de thèse, Charlotte Delmas a en particulier finalisé, et complété un travail poursuivi en interne chez GE Healthcare de segmentation du micro-guide dans les images fluoroscopiques [Del17, chapitre 2]. L’algorithme développé permet d’extraire le micro-guide sous forme de squelette indépendamment pour chaque image et par conséquent sans faire appel à une régularisation temporelle. Les différentes étapes de l’algorithme sont illustrées sur la Figure 3.19. Un exemple de résultat de segmentation est présenté dans la Figure 3.20. Elle illustre les propriétés caractéristiques de ce pré-traitement : une segmentation globalement de bonne qualité, avec cependant quelques faux positifs (barbules, structures secondaires), et quelques faux négatifs (trous dans la segmentation).

C’est donc avec cet a priori technique fort (disponibilité d’une segmentation du guide sous forme de squelette en 2D) que nous avons abordé, dans le cadre de la thèse de Charlotte Delmas, la question de la reconstruction stéréoscopique du micro-guide. Notre méthode a été publiée dans [15] (voir aussi [Del17, chapitre 4]).

3.5.2 Reconstruction stéréo de courbes

La reconstruction stéréoscopique d’une scène peut se concevoir sous une forme dense, où une disparité est calculée pour chaque pixel des images [SS02], ou sous une forme plus éparse où seules certaines structures, supposées caractéristiques de la scène, sont reconstruites : points, courbes

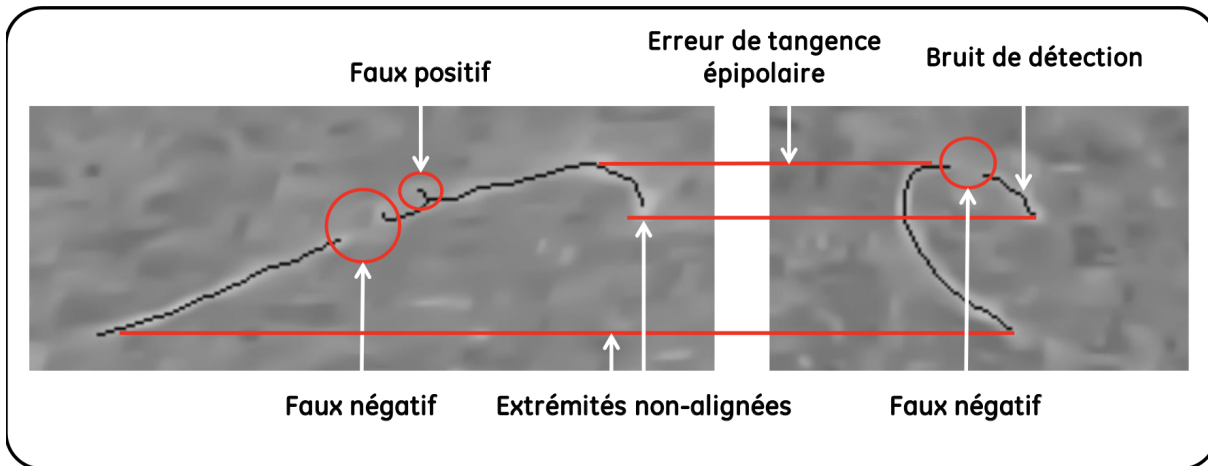


FIGURE 3.20 – Exemple de segmentation du micro-guide dans une paire d’images rectifiées (les lignes épipolaires sont horizontales). Malgré une bonne qualité générale, le résultat peut présenter des trous (faux négatifs), des barbules ou fragments parasites (faux positifs), et des imprécisions qui, combinées aux incertitudes de calibration peuvent poser des problèmes aux extrémités ou aux zones de tangence épipolaire.

(contours), portions de plans voire un modèle de forme et/ou d’apparence [YJS06]. Dans notre cas, l’objet à reconstruire est connu, une courbe unique, ouverte et se déformant de manière non-rigide tout en restant lisse.

Un corpus bibliographique assez abondant s’est penché sur le problème spécifique de la reconstruction de courbes, que ce soit des segments de droite [CKP95, HS89] ou des courbes plus générales qui permettent de modéliser le contour des objets [RF91, BmH01, KA03, MGBL07, FK10] ou des structures arborescentes [CVMR02, LZ06]. Le problème de la stéréovision peut se décomposer en trois sous-problèmes [BBH03] : la *calibration*, comme «procédé de détermination de la géométrie interne et externe des caméras», la *correspondance* comme «détermination des endroits où un même point physique se projette dans les caméras» et la *reconstruction*, qui consiste à «déterminer la structure tri-dimensionnelle.»

La calibration nous est donnée par le recalage basé machine décrit plus haut. Nous avons de plus supposé dans ce travail que cette calibration ne comportait pas d’erreur. La Figure 3.21, où la paire d’images est rectifiée, permet une première réflexion sur les questions de correspondance et reconstruction de courbes sur un cas, que nous nommerons *cas trivial*. La correspondance peut se concevoir entre fragments de courbe (ici, un seul par image, couvrant la même bande épipolaire) ou au niveau de chaque point des fragments (correspondance univoque dans ce cas aussi puisque chaque ligne épipolaire coupe chaque fragment en au plus un point). Les vues étant parfaitement calibrées, la reconstruction de chaque point de la courbe 3D peut se faire par triangulation, et la structure de courbe se déduit du paramétrage dit *épipolaire* induit par l’indice de ligne de chaque point d’un fragment (les deux images génèrent le même paramétrage épipolaire), et qui permet de chaîner les points 3D en courbe.

Le problème de mise en correspondance, ou *appariement*, appelé aussi *problème d’ouverture* pour les courbes [BmH01], est commun au suivi visuel où il est également nommé *association de données*. Les correspondants potentiels d’un point sont recherchés soit le long de la ligne épipolaire dans l’autre image [BvdKvW⁺03, LZ06], soit le long de la normale à la courbe estimée [MGBL07, WHM⁺15], soit de manière implicite en suivant les géodésiques d’un champ de force [CVMR02]. Ne pas se fier à la géométrie épipolaire permet de gérer les imprécisions de calibration mais nécessite une bonne estimation de la courbe, soit par transport temporel, soit par initialisation manuelle [CVMR02]. Les candidats sont sélectionnés, ou pondérés, sur la base de critères photométriques basés sur le gradient de l’image [BMP⁺99, KA03, MGBL07], ou une similarité de l’apparence locale [SZ00, WHM⁺15]. Une similarité de forme entre courbes peut également être employée [GW99], mais elle n’est envisageable dans le contexte stéréo que si

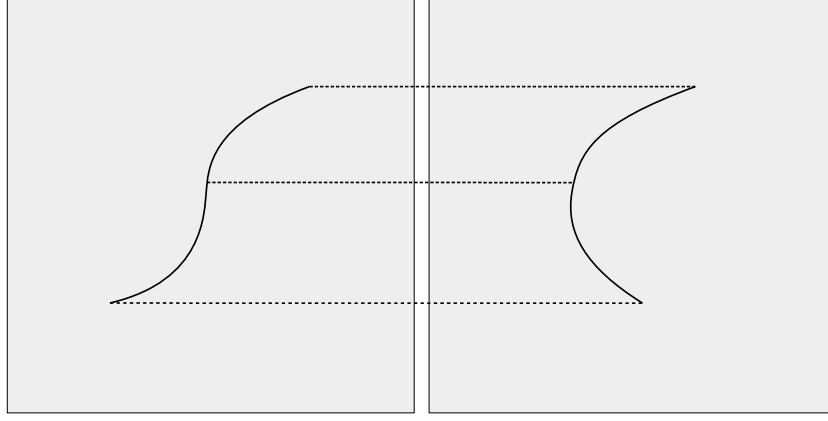


FIGURE 3.21 – Cas trivial de reconstruction de courbes à partir de deux vues calibrées et segmentées (lignes pleines). Les images sont présentées en géométrie rectifiée : les correspondances épipolaires sont établies par les lignes horizontales dont trois exemples sont donnés en pointillés (extrémités et une position intermédiaire). Cette correspondance est univoque pour chaque point des courbes, rendant triviale la reconstruction par triangulation.

les vues sont proches. Des contraintes géométriques classiques peuvent aussi s'appliquer, comme la limite de disparité ou des contraintes d'ordre mais ces critères n'ont pas d'impact face à la contrainte de continuité figurale [SZ00] et peuvent être violées dès que les vues sont éloignées ou si la structure à reconstruire s'auto-occulte [LZ06]. Le découpage épipolaire [FK10] permet de structurer les courbes présentes dans la segmentation en fragments de courbe qui suivent le cas trivial : fragments connexes ne présentant qu'une intersection au plus avec toute ligne épipolaire. Chaque hypothèse d'appariement équivaut dès lors à un fragment de courbe 3D qui peut être validé par reprojection dans d'éventuelles vues supplémentaires [KA03, FK10] ou plus localement par transfert épipolaire [SZ00].

La reconstruction peut se passer de mise en correspondance (méthodes « *correspondenceless* » [BBH03]) en cherchant à ajuster un modèle de courbe 3D aux images [CVMR02, MGBL07, BDD09]. On peut cependant considérer que cette procédure est aussi un moyen d'établir une correspondance par recalage et les deux étapes d'appariement et de reconstruction sont souvent avantageusement couplées [KA03]. Dans le cas d'images segmentées, un cas très fréquent est celui où les courbes 2D sont paramétrées, par exemple par chaînage de pixels ou par ajustement d'une spline 2D. Une paramétrisation de la courbe en 3D peut être établie à partir de cette paramétrisation, sous contrainte de continuité locale [BvdKvW⁺03] ou plus globale [HBJ⁺12], voire une régularité locale d'ordre supérieur [LZ06], ou encore par un critère global de similarité affine [BmH01]. Sans paramétrisation 2D, une alternative consiste à reconstruire un point 3D par hypothèse d'appariement ponctuel. La reconstruction consiste alors à extraire une courbe de ce nuage de points 3D, par exemple en chaînant ces points 3D via un critère de distance euclidienne [Lee00, PLDH11] ou de cohérence structurelle [Lin03].

La reconstruction peut mener à un ensemble de fragments de courbe en 3D. Un cas typique est celui où un découpage épipolaire permet de se ramener au cas trivial pour chaque paire de fragments 2D [FK10]. Une deuxième étape de reconstruction peut alors être menée au niveau de ces fragments, que l'on va chercher à regrouper et ordonner en structures courbes. Si l'espace entre deux fragments 3D n'est pas trop important, la qualité du transport du repère local de Frenet aux extrémités offre un bon critère de connexion [LZ06]. L'approximation locale par *curvelets* [FK10] ou plus globale par B-spline d'un couple de fragments [KA03] permet aussi de valider la connexion en évaluant la qualité d'un support image sur la partie interpolée.

3.5.3 Principe de l'algorithme proposé

Dans notre cas, le guide est distingué du fond par un critère photométrique dans l'étape de segmentation, mais la photométrie du guide est constante dans les images et elle ne peut donc

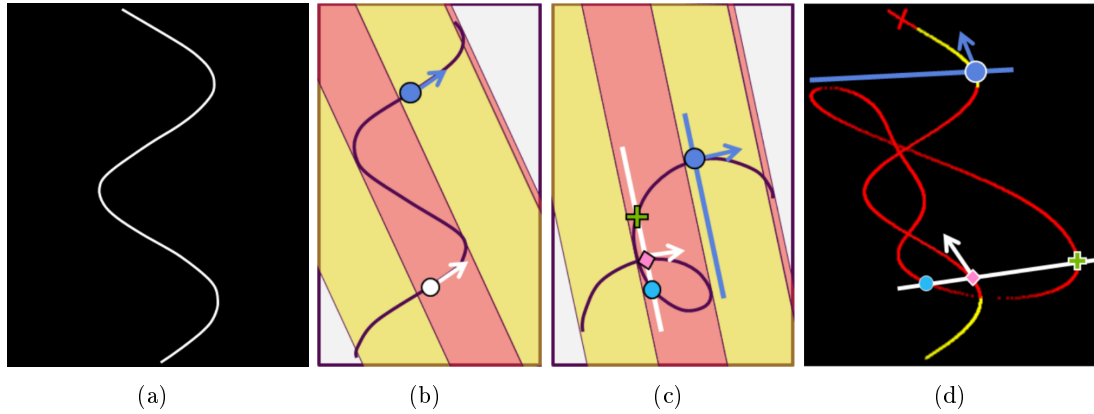


FIGURE 3.22 – Exemple de reconstruction épipolaire : (a) courbe initiale (une hélice) ; (b) première vue en projection ; (c) deuxième vue en projection ; (d) reconstruction épipolaire issue de ces deux vues. Le découpage épipolaire apparaît dans les vues en projection en alternant les couleurs de fond des bandes épipolaires (jaune/rose). Chaque bande épipolaire correspond : soit au cas trivial (cas des bandes jaunes), auquel cas un seul fragment 3D, dit univoque, peut être reconstruit (fragments jaunes dans la reconstruction épipolaire, l'unicité de la correspondance épipolaire apparaît avec l'exemple des points bleus) ; soit plusieurs fragments 3D peuvent être reconstruits (le point blanc dans la première vue a trois correspondants potentiels dans la deuxième vue, ce qui donne lieu à trois hypothèses de reconstruction). Cette ambiguïté est responsable de la complexité de la reconstruction épipolaire. La difficulté consiste à sélectionner et ordonner les bons fragments 3D dans ces zones.

pas aider dans la phase d'appariement. Nous avons donc exploité la contrainte épipolaire, sans pondération des différents candidats éventuels à l'appariement.

Dans l'hypothèse où une seule courbe algébrique est présente dans les images, la reconstruction stéréoscopique admet deux solutions, dont une est la courbe recherchée. Leur union est une variété formée par l'intersection des cônes s'appuyant sur les courbes projetées et ayant pour sommet les centres optiques des caméras [KS04]. Si on considère la segmentation dans les images comme un ensemble de pixels, le nuage de points 3D obtenu par triangulation à partir de toutes les hypothèses d'appariement épipolaire, est un échantillonnage de cette variété. Un découpage épipolaire, suivi de la reconstruction de tous les appariements épipolaires des fragments 2D, permet de donner une première structure à ce nuage sous la forme d'un ensemble de fragments 3D. Cependant, même en l'absence de toute erreur de calibration ou de bruit de segmentation, la forme reconstruite reste complexe. La Figure 3.22 montre un exemple d'une telle reconstruction que nous nommons *reconstruction épipolaire*. La courbe recherchée étant un sous-ensemble de la reconstruction épipolaire, notre algorithme va sélectionner et ordonner un sous-ensemble des fragments 3D, que nous appelons *support*, puis la courbe sera reconstruite par approximation B-spline cubique de ce support.

3.5.4 Pré-sélection des fragments 3D

Nous pouvons poser le problème sous forme de graphe où chaque fragment 3D est un nœud, une arête signifie que les deux fragments qu'elle relie sont connectés et son poids mesure la qualité de la connexion. Tout chemin de ce graphe est un support potentiel.

Cependant, en l'absence d'a priori supplémentaire, le graphe est initialement complet, ce qui rend la combinatoire du problème insurmontable (voir [Del17, chapitre 4] pour un exemple détaillé). Nous posons donc quatre conditions :

1. Nous ne considérons que les fragments 3D de longueur minimale 2 voxels. Cette idée que les fragments 3D plus courts sont moins fiables se retrouve aussi chez [KA03, FK10]. Mais nous rejetons aussi les fragments 3D qui s'étendent en reprojection sur moins de trois

lignes épipolaires dans une vue⁶ : ces fragments sont en effet dans, ou proches d'un plan épipolaire et leur reconstruction est donc imprécise.

2. Après découpage épipolaire, un cas particulier est celui où un seul fragment est visible au niveau d'une bande épipolaire dans chacune des deux images. C'est le cas trivial et un seul fragment 3D peut être reconstruit. Nous considérons que la présence simultanée d'un seul fragment dans les vues est une confirmation de la validité de la segmentation d'une vue par l'autre vue. Par conséquent, nous faisons l'hypothèse que le fragment reconstruit, dit alors *univoque* fait partie du support recherché. De tels fragments vont servir de points d'amer à la reconstruction.
3. Nous recherchons à reconstruire une courbe simple et régulière en 3D. Par ailleurs, dans le contexte de navigation vasculaire, cette courbe a globalement la même allure que la ligne centrale des vaisseaux sanguins. Nous faisons donc l'hypothèse qu'il n'y a pas ou peu d'auto-occultation de la courbe. Ce critère d'unicité est également invoqué dans [KA03].
4. Enfin, nous recherchons un support de longueur maximale, une fois remplies les trois conditions précédentes.

Le premier critère a pour conséquence une réduction du nombre de nœuds dans le graphe. Ces nœuds sont des petits fragments 3D qui n'auront que peu d'impact sur la forme de la courbe, mais peuvent aussi être des fragments plus longs situés aux zones de tangence épipolaire ou aux extrémités. Le critère de connexion des fragments devra résister à ces trous éventuellement grands.

Le deuxième critère produit une réduction drastique du nombre d'arêtes dans le graphe. Comme il oblige tout support à contenir les fragments univoques, la continuité figurale impose qu'un fragment ne peut être connecté à un autre fragment si leurs projections sont séparées par une bande épipolaire donnant un fragment univoque. Toutes les arêtes correspondant à cette situation sont éliminées du graphe. La recherche du support global est alors subdivisée en recherches de supports dans les zones intermédiaires entre deux fragments univoques ou aux extrémités. Cette subdivision est illustrée sur la Figure 3.23 : elle se traduit directement sur le graphe initial, qui peut alors se découper en sous-graphes dans chacun desquels un chemin est recherché. L'objectif est que ce chemin se connecte bien avec le(s) fragment(s) univoque(s) adjacent(s). Il est à noter que le graphe n'est plus complet mais chaque sous-graphe le reste. Par ailleurs, pour chaque sous-graphe, la reprojection des fragments 3D le formant s'étend sur une ou plusieurs bandes épipolaires adjacentes.

Le troisième critère agit également sur les arêtes, au niveau des sous-graphes. L'absence d'auto-occultation se traduit par l'impossibilité pour deux fragments 3D d'un même support de se projeter sur le même fragment 2D. Ne pouvant appartenir au même support, ils ne peuvent être connectés. Cette interdiction induit ainsi la disparition du graphe de toute arête reliant deux fragments 3D qui se reprojettent sur le même fragment 2D. Il est important de noter que cette condition ne s'applique qu'à deux fragments 3D issus d'une même bande épipolaire.

À ce stade, et malgré la simplification du graphe initial, la combinatoire reste encore trop importante. La sélection des fragments est rendue plus efficace en remarquant tout d'abord qu'après avoir appliqué la deuxième condition (fragments univoques), chaque sous-graphe est complet. L'application de la troisième condition leur fait perdre cette qualité mais les seules arêtes retirées sont celles qui relient deux fragments incompatibles, c'est-à-dire ne pouvant apparaître dans le même support. A contrario, un support admissible correspond alors forcément à une clique du sous-graphe considéré. **Le quatrième critère**, qui veut que nous recherchions les supports maximaux, conduit à ne sélectionner que les ensembles de fragments formant des cliques maximales de chaque sous-graphe.

L'énumération des cliques maximales d'un sous-graphe peut s'avérer très lourde en temps de calcul (problème NP-difficile). Mais nous avons remarqué plus haut que le troisième critère n'agit qu'au niveau de chacune des bandes épipolaires que couvre le sous-graphe en projection. Pour un sous-graphe donné, nous commençons donc par énumérer les cliques maximales du (sous-)

6. Nous travaillons sur des images rectifiées : une ligne épipolaire=une ligne de l'image

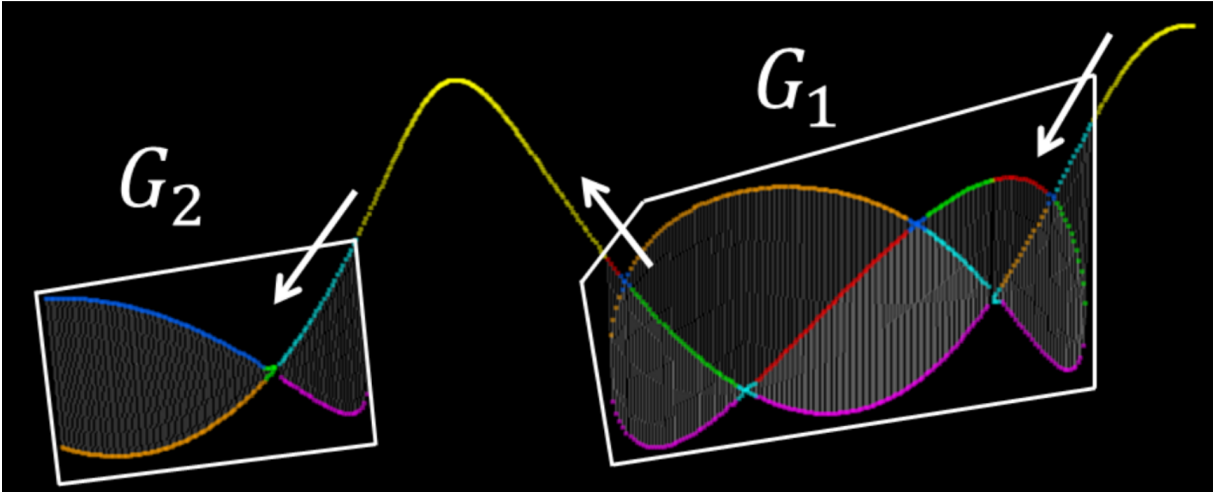


FIGURE 3.23 – La reconstruction épipolaire d’une courbe contient des fragments univoques (en jaune), alternant avec des ensembles de fragments dont seuls certains appartiennent au support de la courbe. Ces dernières zones, intermédiaires, de la reconstruction épipolaire correspondent à deux sous-graphes complets (G_1 et G_2) du graphe initial.

sous-graphe associé à chaque bande épipolaire. Puis nous formons le produit cartésien de ces ensembles. Ce produit est l’ensemble des cliques maximales recherché. Cette recherche peut ainsi être parallélisée suivant les bandes épipolaires et prend en pratique très peu de temps.

Au niveau d’un sous-graphe, l’ensemble des supports candidats est donné par l’ensemble des chemins hamiltoniens de toutes les cliques maximales ainsi énumérées. Dans un premier temps, nous réduisons le nombre de candidats en ne conservant, pour chaque clique maximale, que le chemin hamiltonien produisant le support le plus régulier. Cette recherche passe par la pondération de chaque arête par un coût de connexion afin de la ramener à une recherche de plus court chemin hamiltonien. Dans un deuxième temps, le support global sera bâti en connectant ces candidats locaux aux fragments univoques adjacents. Le support le plus régulier sera retenu comme solution.

3.5.5 Ordonnancement des fragments 3D en supports candidats

La régularité d’une courbe se traduit chez [LZ06] par un bon transport du repère de Frénet, le transport idéal suivant le tracé d’un cercle entre les points de la courbe considérés. Une interpolation par *Moving Least Squares* [Lee00] ne peut combler que de petits espaces. Des interpolations au moyen d’une courbe algébrique sont coûteuses [KA03, FK10] ou sensibles à l’estimation fiable d’informations différentielles locales [TK07]

Nous définissons un support régulier comme étant continu (continuité des positions ou C^0) et lisse (continuité des tangentes ou C^1). La discontinuité C^0 entre deux fragments est mesurée par la distance euclidienne d entre les deux extrémités reliées. La discontinuité C^1 est évaluée par l’angle maximum β observé à chaque extrémité entre la tangente au fragment et le segment qui relie les deux extrémités. La Figure 3.24 donne les éléments de calcul de ces deux valeurs d et β . Nous pouvons ainsi définir le coût de connexion entre deux fragments 3D par :

$$w = \alpha \left(1 - \exp \left(-\frac{d^2}{\sigma^2} \right) \right) + (1 - \alpha) \left(\frac{1}{2} (1 - \cos(\beta)) \right) \quad (3.21)$$

Le premier terme pénalise les grandes distances, avec un effet de seuil doux en-dessous de σ , et le deuxième terme pénalise les fortes courbures. α permet de pondérer ces deux termes.

Une arête dans le graphe exprime la possibilité de connecter deux fragments 3D. Cette possibilité se décline sous quatre versions (une par association possible des deux extrémités de chaque fragment). Nous pondérons ainsi chaque arête avec 4 coûts de connexion associés. La recherche

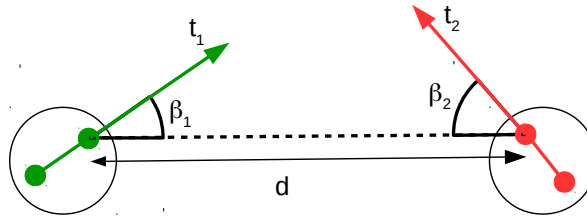


FIGURE 3.24 – Le caractère C^0 d’une connexion entre deux fragments 3D (vert et rouge, seules les portions extrémales sont montrées) est indiqué à travers la distance entre les deux extrémités connectées (longueur de la ligne en pointillés). Le caractère C^1 est indiqué à travers l’angle $\beta = \max(\beta_1, \beta_2)$. t_1 (resp. t_2) est la tangente en l’extrémité connectée du premier (resp. deuxième) fragment 3D. Cette tangente est évaluée par régression linéaire sur un voisinage de l’extrémité considérée (cercles noirs).

du chemin le plus régulier se traduit ainsi par la recherche du chemin hamiltonien de coût minimum. C’est le problème du Voyageur de Commerce. Nous avons donc suivi une approche au plus proche voisin en prenant successivement pour hypothèse de départ chaque fragment du graphe, puis en construisant le chemin hamiltonien de coût minimal en ajoutant itérativement le voisin le plus proche. Cette heuristique permet d’inclure aisément la condition qu’une extrémité de fragment ne peut être connectée qu’une seule fois. Son exécution sollicite la majorité du temps de calcul de notre algorithme.

Tous les chemins trouvés sont conservés comme supports candidats pour chaque sous-graphe (un chemin par fragment du sous-graphe). Tous les supports obtenus par combinaison d’un support par sous-graphe et des fragments univoques sont qualifiés en calculant un coût de connexion global également basé sur l’équation 3.21 et le support de coût minimal est conservé. La courbe finale est obtenue par approximation B-spline cubique de la suite de fragments formant ce support.

3.5.6 Résultats et discussion

Les résultats détaillés sont disponibles dans [Del17, chapitre 4]. Nous n’en reprenons ici que les conclusions principales.

Données simulées

Afin de disposer de trajets de navigation à la fois réalistes et tortueux, nous avons extrait, dans la 3DRA chez deux patients, la ligne centrale de 12 vaisseaux sanguins afférents à une malformation artério-veineuse (6 par patient). 2289 courbes simulant l’extrémité distale d’un micro-guide ont été formées en échantillonnant chaque ligne centrale sur une longueur de 2 cm, tous les millimètres. Deux vues ont été générées pour chaque courbe grâce à deux matrices de projection issues de la calibration d’un système biplan. Parmi ces courbes, 953 correspondaient au cas trivial. L’évaluation a porté sur les 1336 restantes.

Cette étude nous a permis de déterminer expérimentalement un paramétrage efficace pour notre algorithme. Par ailleurs, 97% des supports reconstruits étaient corrects, avec une distance moyenne à la courbe vérité terrain de moins de 0.2 mm (taille d’un voxel en 3DRA). L’erreur de reconstruction du point extrémal de la courbe est en moyenne du même ordre, mais avec un écart-type plus élevé (0.4 mm).

Données fantômes

Nous avons utilisé un fantôme d’arbre vasculaire cérébral creusé dans un bloc de silicone (Elastrat, Genève). Un micro-guide avec une extrémité distale légèrement courbée et de longueur radio-opaque 2 cm a été inséré dans le fantôme à l’aide d’un cathéter porteur. Une séquence de fluoroscopie en biplan a été acquise, suivie d’une acquisition 3DRA, sur un système vasculaire

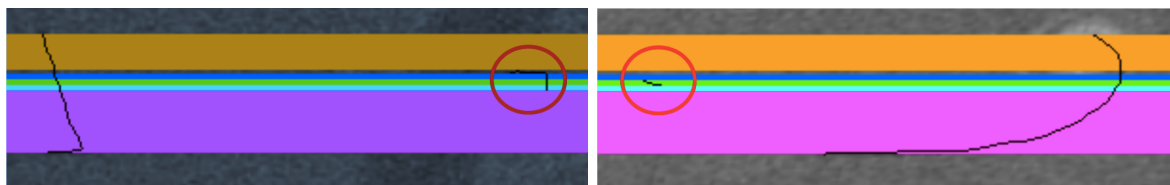


FIGURE 3.25 – Échec de la reconstruction due à des faux positifs (données sur fantôme, les couleurs indiquent les bandes épipolaires). Les erreurs de détection sont entourées en rouge dans les vues. Étant situées au niveau de la même bande épipolaire (verte), elles génèrent un deuxième fragment 3D dans le sous-graphe de cette bande, qui, sans elles, n'aurait contenu qu'un fragment univoque. Aucun de nos critères ne permet d'éliminer le fragment erroné.

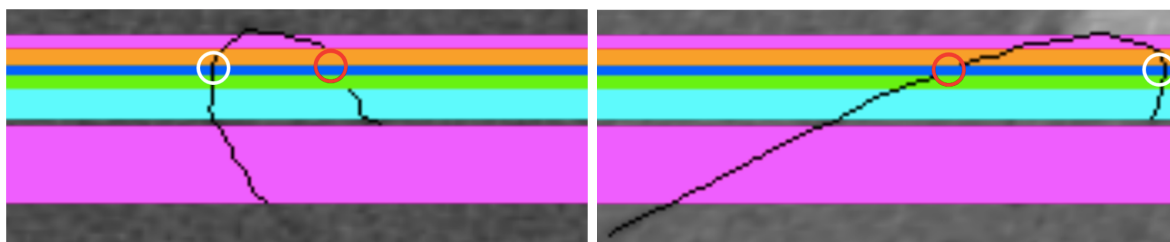


FIGURE 3.26 – Échec de la reconstruction dû à des faux négatifs (données sur fantôme, les couleurs indiquent les bandes épipolaires). Les deux fausses détections sont entourées en rouge dans les vues. Elles sont situées au niveau de la même bande épipolaire (bleue). Cette bande contient un seul vrai positif dans chaque vue (entourés en blanc). Le sous-graphe associé à cette bande ne contient donc qu'un fragment univoque. Suivant le deuxième critère, ce fragment est inséré dans la solution alors qu'il est erroné.

biplan Innova 3131IQ (GE Healthcare), pour 4 positions différentes du micro-guide immobile (30 paires d'images par position). Une segmentation suivie d'une reconstruction a été faite pour chaque paire d'images de toutes les séquences. La ligne centrale du micro-guide a été extraite de la 3DRA pour servir de vérité terrain.

L'objectif de cette expérience était d'évaluer l'impact du bruit dans les images fluoroscopiques sur la chaîne de traitement menant à la reconstruction.

L'erreur moyenne de reconstruction le long de la courbe est de 0.54 mm avec un écart-type de 0.22 mm. L'erreur de reconstruction de l'extrémité est en revanche plus forte (de 0.56 mm à 2.1 mm selon le cas). Dans quelques cas, la reconstruction a échoué, suite aux erreurs de segmentation. Dans un cas (voir Figure 3.25), la présence d'un faux positif dans les deux vues au niveau d'une même bande épipolaire a généré un fragment 3D erroné intégré à toutes les cliques maximales de cette bande, et donc au support final. Dans un autre cas (voir Figure 3.26), la présence d'un faux négatif au niveau d'une même bande épipolaire, mais ne concernant pas le même endroit de la courbe dans les deux vues, a généré un fragment univoque, donc intégré à la solution, mais erroné.

Données cliniques

Nous avons collecté 7 séquences d'acquisition biplan d'un micro-guide en mouvement pour 4 patients différents donnant lieu à un total de 272 reconstructions. Aucune vérité terrain n'est disponible pour les données patients. Nous avons donc comparé la reprojection de notre courbe reconstruite à une segmentation semi-manuelle du guide dans les images fluoroscopiques.

En l'absence de structure perturbant fortement la segmentation (par exemple des coils placés lors d'une précédente intervention), l'erreur moyenne le long de la courbe reprojetée est de 2 pixels (écart-type de 1 pixel, le pixel vaut 0.16 ou 0.2 mm en fonction des paramètres de l'acquisition). Cette erreur monte cependant à 6.4 pixels (écart-type de 8.6 pixels) quand la segmentation comporte trop de fausses détections.

Quelques tentatives de reconstruire le coil (également une courbe opaque aux rayons X) ont montré que notre méthode ne fonctionne plus au-delà du déploiement des premières boucles : la combinatoire redevient rédhibitoire, mais de plus, le troisième critère (absence auto-occultation) est rapidement rendu caduque.

3.6 Conclusion et perspectives

Dans les premières années suivant ma thèse, la 3DRA était largement méconnue des chercheurs en imagerie médicale qui la percevaient comme redondante avec l'Angio-CT ou -IRM, avec la restriction supplémentaire de sa seule utilité à la neuroradiologie interventionnelle. Nos contributions, faisant écho à d'autres par des chercheurs en imagerie médicale ou des cliniciens au début des années 2000, ont aidé à en préciser l'intérêt unique dans le traitement des anévrismes et des MAV intracrâniennes, à mieux en cerner la spécificité par rapport aux autres modalités d'imagerie et à susciter des travaux visant à en étendre le principe suivant diverses directions de recherche.

Tout d'abord, la 3DRA a rapidement été rendue disponible sur les salles vasculaires, pour le cathétérisme des artères du cou, de l'abdomen et des membres inférieurs dont le mouvement peut être réduit en arrêtant momentanément la respiration. Une 3DRA peut également être acquise en cardiologie interventionnelle, mais elle requiert en général une synchronisation ECG, ce qui allonge le temps d'acquisition et le volume de produit de contraste [LLSK10]. Une ligne de travaux cherche donc à exploiter une acquisition habituelle (rotation sur 5 secondes), avec l'espoir supplémentaire de capturer la déformation des artères en 3D lors d'un battement cardiaque [BVMA04].

Une autre voie consiste à transposer l'acquisition rotationnelle sur un arceau mobile chirurgical, en vue de produire une reconstruction tomographique ou pour des applications de réalité augmentée. La calibration d'un tel système n'est cependant pas aisée car ces arceaux n'ont pas d'ancrage fixe dans la salle d'opération. Les prototypes décrits dans la littérature emploient un capteur de localisation [OAA⁺12] ou encore une caméra [NHT10, FBW⁺16] pour localiser l'arceau par rapport au patient. Pour notre part, et malgré l'intérêt scientifique des publications liées, nous ne croyons pas en l'avenir industriel et clinique de tels systèmes. Les fonctionnalités avancées que sont la reconstruction et la réalité augmentée ne seront disponibles que sur des systèmes haut de gamme ciblant les équipes spécialistes, et les cas cliniques difficiles qu'elles traitent. Cet environnement clinique et industriel sera plus favorable à des systèmes plus évolués comme le Zeego de Siemens, qui, tout en étant plus léger et ouvert qu'un arceau vasculaire, est un arceau chirurgical fixé au sol. Dans ce cas, les méthodes de calibration que nous avons développées sont tout à fait applicables et ont prouvé leur efficacité.

Un troisième axe d'extension est celui de la multimodalité. Le concept de système hybride a été exploré dans le milieu des années 2000 [RSB⁺05, BGP⁺06]. Ces salles font cohabiter un arceau vasculaire avec un scanner IRM qui partagent la même table d'opération. Ce concept est aujourd'hui tombé en désuétude, ce que nous relierons directement à la capacité actuelle des arceaux rayons X de produire des reconstructions tridimensionnelles. Dès notre thèse, nous avons considéré que la 3DRA, étant à la fois acquise en salle d'opération et tridimensionnelle, constituait un pont idéal entre les images per-opératoires 2D et des modalités péri-opératoires 3D et en premier lieu l'IRM [38]. Le rôle central que peut jouer la 3DRA est très bien repris dans [RHM⁺11]. La qualité actuelle des reconstructions 3D acquises par arceau est un signe de la convergence qui s'opère depuis une quinzaine d'années avec les scanners CT classiques⁷. Le problème multimodal se ramène donc au problème du recalage CT/IRM, bien connu, maîtrisé, et pour lequel des solutions commerciales existent.

Les perspectives de la Fluoroscopie Augmentée sont, à notre avis, d'ordre industriel et clinique où il s'agit d'en améliorer l'ergonomie et parfaire son insertion dans la pratique quotidienne du neuroradiologue interventionnel. D'un point de vue technique, les défis de la réalité augmentée en imagerie médicale concernent la gestion de la déformation [7, 18, 19, 22], et l'ergonomie de la

7. Tous ces systèmes sont aujourd'hui multi-barrettes, autrement dit utilisent un capteur plan.

visualisation. Le plus gros défi, qui bénéficiera en premier lieu à la neuroradiologie interventionnelle, nous paraît cependant de passer du 2D au 3D, c'est-à-dire proposer une reconstruction 3D, et donc une visualisation 3D, de la scène chirurgicale en temps réel. Les travaux exposés dans ce chapitre n'en sont que les prémices.

L'algorithme décrit plus haut ne permet pas de rejeter certains fragments, notamment les fragments univoques erronés. Un retour à l'image nous semble la meilleure stratégie pour invalider ces fragments [FK10]. Une contrainte de cohérence temporelle permettrait de limiter les variations de longueur, de maintenir une zone d'intérêt contenant la reconstruction, et permettrait de paramétrer en 2D des fragments plus longs que les fragments épipolaires, notamment pour éviter les trous artificiellement formés aux zones de tangence épipolaire. Conserver ces zones de tangence est selon nous essentiel pour prendre en compte l'incertitude de la calibration et développer un algorithme proposant simultanément la reconstruction du micro-outil et un affinage de la calibration [MPLŠ15, NF15].

Le plus grand défi est cependant, à notre avis, celui de la reconstruction monoculaire [vWBN05, BGK⁺08, GBK⁺09, BDD09]. Nos travaux en reconstruction stéréo sont une étape essentielle afin de disposer d'une méthode de référence pour valider une reconstruction dynamique [17]. Mais ils reposent sur une acquisition simultanée par les deux arceaux frontal et latéral, ce qui n'est pas possible stricto sensu⁸ et ne correspond pas à la pratique médicale. Les approches actuelles reposent sur des modèles de formes que nous pensons inadaptés aux brusques changements que peut connaître la forme d'un micro-outil. Nous suivons plutôt une approche de simulation guidée par l'image, qui incorpore des modèles dynamiques à la fois plus complexes et plus réalistes. Mais cela nécessite de développer une simulation physique du déploiement des micro-outils en temps interactif. C'est l'objectif des travaux exposés dans le chapitre suivant.

8. La fluoroscopie, pulsée, est alternée entre arceaux frontal et latéral pour limiter le fond diffus.

Chapitre 4

Contributions à la simulation interactive pour la neuroradiologie interventionnelle

Sommaire

4.1	Introduction	81
4.1.1	Contexte	81
4.1.2	Contributions	84
4.2	Segmentation vasculaire : l'algorithme RBT	85
4.2.1	Description générale de l'algorithme	85
4.2.2	Extraction des points candidats	85
4.2.3	Ajustement de cylindre par RANSAC	86
4.2.4	Construction de l'arbre	87
4.2.5	Validation	89
4.3	Reconstruction vasculaire : l'algorithme LIM	93
4.3.1	Modèle implicite	93
4.3.2	Formulation énergétique	96
4.3.3	Optimisation	97
4.3.4	Validation	99
4.4	Application à la simulation : implantation et validation	101
4.4.1	Contexte de la simulation	102
4.4.2	Implantation de notre modèle dans un contexte simulation	102
4.4.3	Quelques éléments pour la validation d'une simulation interactive	107
4.5	Conclusion et perspectives	110

4.1 Introduction

4.1.1 Contexte

Au milieu des années 1990, les progrès réalisés dans le domaine de la simulation physique par ordinateur avaient permis la construction de systèmes de réalité virtuelle de très grande qualité pour les industries de l'aviation ou de la défense. Leur réalisme était tel qu'ils avaient un impact positif sur la formation ou la certification du personnel ainsi que sur le design et le test de nouveaux matériels. L'extension de ce principe à la chirurgie, qui partage ces grandes problématiques, était donc naturelle, d'autant plus que cette période a également été celle de la maturation des techniques de chirurgie minimalement invasive [DK98]. Cet ensemble de pratiques présente en effet deux caractéristiques qui les rendent particulièrement compatibles avec un environnement de simulation : le geste se fait grâce à un outil dont la forme géométrique et le comportement mécanique sont a priori plus aisément modélisables que la main, et ce geste est contrôlé à distance par l'intermédiaire d'un système d'imagerie visualisé sur écran.

Simuler la radiologie interventionnelle

La radiologie interventionnelle par cathétérisme vasculaire a été parmi les premières spécialités à bénéficier de prototypes de simulateurs. Bien que d'autres systèmes aient été proposés (comme le système *CathI/HERMES* [RBSH04, AdPP06]), deux grandes lignées de simulateurs ont émergé des systèmes pionniers *da Vinci* [ABK⁺96, AR98] et *Dawson-Kaufman* [HMM⁺95, Meg96]. Initialement conçus comme des plateformes génériques, ils se sont spécialisés en cardiologie vasculaire avec *ICard* [WCL⁺98] basé sur *da Vinci* et *ICTS*¹ [CDM⁺00] sur *Dawson-Kaufman*, et en neuroradiologie interventionnelle avec *NeuroCath* [NC01, Ma07] issu de *da Vinci*. Ce sont de véritables environnements simulés qui sont proposés d'emblée et non pas juste des logiciels. Cette démarche a permis de progresser suivant trois directions en parallèle. En premier lieu, des systèmes commerciaux ont pu être développés et aujourd'hui les systèmes CathLabVR², ANGIOmentor³, VIST⁴, et COMPASS⁵ implantent des simulateurs de procédures de radiologie interventionnelle par cathétérisme vasculaire. Par ailleurs, la prise en main de ces environnements par des médecins a permis d'approfondir la compréhension de leur utilité, et de préciser les conditions de leur impact sur la pratique des médecins en termes de formation au geste chirurgical et de planification opératoire [ACC⁺02, Daw06, Gou10, GCJ⁺12], si bien que le potentiel en est reconnu aujourd'hui dans la littérature médicale [PD17] et que l'utilisation de la simulation est recommandée dans le cursus des neuroradiologues interventionnels [PRBT17]. Enfin, quatre grands défis technologiques ont pu être identifiés : retour haptique, visualisation, modélisation physique, modélisation géométrique. Les progrès réalisés dans ces domaines ont été progressivement intégrés aux systèmes existants, par exemple *NeuroCath* [CCY⁺03] ou *ICTS* devenu *EVE*⁶ [WPL⁺05] puis *SOFAEVE* lors de son intégration dans la plateforme SOFA⁷ développée par Inria. La modélisation physique a en particulier connu de nombreuses avancées, pour modéliser le flux vasculaire [WAC07], mais aussi et surtout pour modéliser le comportement du cathéter en temps interactif, par des modèles à éléments finis non-linéaires permettant de prendre en compte de grandes déformations, et par la gestion des contacts mécaniques entre les outils et la paroi vasculaire [24, 26, DCLN06, DLC07, LHG⁺09].

La modélisation géométrique a en revanche été plus rarement abordée [LG98, DLC07, WLK⁺11]. La nécessité d'une modélisation adaptée à la simulation a pourtant été établie de longue date [Meg96]. Les concepteurs de *NeuroCath* relèvent même qu'une modélisation vasculaire adaptée à la simulation doit intégrer à la fois la géométrie et la topologie de l'arbre vasculaire [NC01]. Le réalisme d'une simulation physique dépend en effet de la qualité des informations de collision [TKH⁺05]. Dans le contexte vasculaire, où l'outil interventionnel simulé est en contact permanent avec la surface des vaisseaux sanguins, la précision et la performance de la simulation dépendront donc fortement de la représentation choisie pour cette surface.

Représenter la surface vasculaire

Trois grandes familles existent pour la représentation d'une surface : polygonale, paramétrique et implicite. Extraire une telle surface de données image est l'acte de segmentation pour lequel il existe une vaste littérature [LAB⁺09]. Dans un objectif de visualisation, on recherche une surface lisse dans sa globalité et assez précise pour préserver visuellement les structures d'intérêt [PO08, WHM13]. Dans un contexte de simulation, les informations de collision sont extraites localement et requièrent une grande précision [TKH⁺05]. Ce point de vue un peu différent a été discuté dans [LG98] mais de nombreuses modélisations vasculaires ont vu le jour depuis cet ancien travail.

-
1. Interventional Cardiology Training System
 2. CAE Healthcare (<http://caehealthcare.com>)
 3. Simbionix (<http://simbionix.com>)
 4. Mentice AB (<http://mentice.com>)
 5. Simsuite (<http://medsimulation.com>)
 6. EndoVascular Embolization simulator
 7. Simulation Open Framework Architecture (<http://sofa-framework.org>) [FDD⁺12]

Les modèles polygonaux sont les plus couramment employés. Il est assez aisé de contraindre leur forme localement [YCM⁺01] et des structures spécialisées, telles que les volumes englobants hiérarchiques (*BVH* pour *Bounding Volume Hierarchies*), ont été développées pour gérer efficacement les collisions entre objets polygonaux. Cependant elles ne sont pas adaptées à la forme arborescente du réseau vasculaire. De plus, la gestion interactive des collisions avec des objets polygonaux déformables reste un défi [TKH⁺05, LYTM06, KHI⁺07, CTM08, LKTS⁺13].

Un modèle paramétrique peut se déduire de la représentation polygonale par interpolation de Catmull Clark [WLK⁺11, FWB04]. Le réseau vasculaire peut aussi se modéliser à partir de primitives graphiques [HPSP01] ainsi que des cylindres généralisés. Outre que ces modélisations reposent sur l'a priori (inexact) de section circulaire pour les vaisseaux, le calcul de la distance à une surface paramétrique, pour estimer la profondeur d'interpénétration, n'est en général pas trivial.

La représentation implicite canonique est donnée par la fonction de distance euclidienne à la surface. Le concept de cylindre généralisé peut être étendu en convoluant la ligne centrale des vaisseaux par un profil 2D de leur section axiale [HLT12, KGPS13]. Une meilleure continuité longitudinale est obtenue par convolution 3D [OP05] mais ces représentations peuvent manquer de détails à moins de complexifier le squelette des vaisseaux sanguins.

Les surfaces implicites procédurales sont particulièrement adaptées à la simulation. Les tests d'inclusion/exclusion se font par simple évaluation de la fonction implicite. Le calcul des forces de friction est plus simple du fait de leur bon degré de continuité. L'intersection avec un segment peut se résoudre rapidement par un algorithme de Newton-Raphson [DLC07]. De plus, des solutions élégantes ont été proposées pour la gestion des collisions [Gas93].

De nombreuses méthodes existent pour reconstruire une surface implicite à partir d'un ensemble de points candidats [BLN⁺13], même en présence de trous dans cet ensemble [OBA⁺03]. La gestion du bruit et des points aberrants est en revanche un problème difficile [MdGD⁺10]. Cependant, les surfaces implicites ont tendance à fusionner quand elles se rapprochent. Cette propriété est à la base de leur capacité à gérer des changements de topologie, mais pose un problème dans le cas vasculaire car les parois des vaisseaux proches ne sont plus préservées [SNB⁺08]. Le contrôle fin de cette propriété est extrêmement difficile, notamment dans la gestion d'objets multiples ou ayant une topologie complexe [GBC⁺13].

Extraire la topologie des vaisseaux

Une représentation commune pour le réseau vasculaire encode sa topologie dans le graphe des lignes centrales, chaque centre sur ces lignes étant augmenté d'une information sur la forme locale du vaisseau [LAB⁺09]. Cette information est par exemple le rayon de la section locale assimilée à un disque [WNV00]. Une meilleure régularité dans la direction longitudinale est assurée en modélisant le réseau vasculaire comme un chemin 4D (les trois coordonnées du centre, augmentées du rayon local) [DC02, LKGD07]. Des alternatives emploient des modèles locaux de cylindres [LCSK⁺04, WC07, WR07, SSM⁺07, FHKP10], des super-ellipsoïdes [TdTF⁺07] ou déduisent la forme locale de l'analyse du hessien de l'image [FDCR01, HOPJ⁺06, CHDO07]. Cependant, la seule information de rayon local ne suffit pas à reconstruire une surface assez détaillée pour la simulation.

L'a priori du modèle permet de filtrer les structures denses non vasculaires, mais il convient de revenir à l'image pour la reconstruction proprement dite de la surface [WNV00]. Les données angiographiques sont par ailleurs bruitées et l'hypothèse locale de circularité de la section est violée en présence de pathologie, aux bifurcations, ou quand un vaisseau s'approche voire touche une structure dense (os, un autre vaisseau ou un anévrisme). Dans la recherche d'un compromis entre un modèle et un terme image, seuls [TdTF⁺07], avec leur approche probabiliste, peuvent prétendre à une certaine robustesse du terme image.

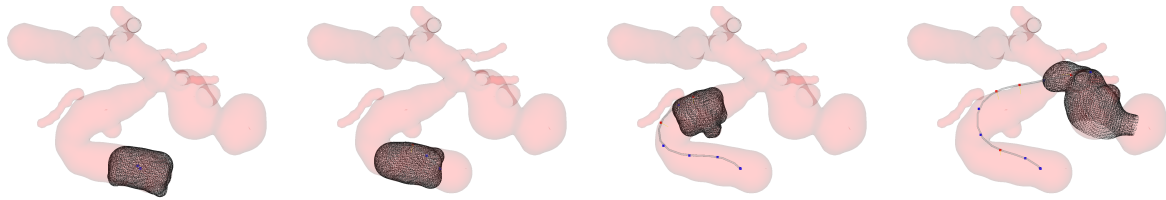


FIGURE 4.1 – Exemple d'utilisation de notre modèle hybride d'arbre de fonctions implicites. Chaque image correspond à un pas de temps de simulation différent. La surface maillée (pour les besoins de la visualisation uniquement) est celle du BM associé à l'extrémité du cathéter. Cette surface implicite locale est utilisée pour résoudre les collisions éventuelles de ce point avec la surface vasculaire.

4.1.2 Contributions

Les travaux repris dans ce chapitre trouvent leur origine dans la rencontre avec Stéphane Cotin, Christian Duriez et Jérémie Dequidt de l'équipe Inria SHACRA de Lille⁸. Cette équipe a mis au point un modèle mécanique pour les micro-outils (micro-cathéter, micro-guide, micro-coil) composé d'une succession de poutres, et qui autorise des temps de simulation interactifs [CDL⁺05, DCLN06]. Des travaux préliminaires avaient montré que de meilleurs temps et une meilleure qualité de simulation étaient atteints avec une modélisation implicite des vaisseaux [DLC07]. Notre rôle, dans cette collaboration, était donc de mettre au point un modèle implicite d'arbre vasculaire.

Suite à des premières investigations prometteuses [Gla09], notre choix s'est porté sur les *Bloppy Models* (BM) [Mur91] pour la modélisation implicite. Mais, outre la consolidation de ces premiers travaux, il fallait remédier à l'absence d'information topologique. Pendant la thèse d'Ahmed Yureidini [Yur14], nous avons ainsi promu un modèle hybride d'arbre de fonctions implicites : l'arbre est celui des lignes centrales des vaisseaux sanguins, et au lieu d'augmenter chaque point d'un rayon, d'un profil de section ou d'un cylindre, nous le faisons par un BM qui modélise la surface du vaisseau sanguin dans un voisinage du point. La Figure 4.1 illustre l'exploitation qui peut être faite d'un tel modèle : pour gérer la collision en chaque point du cathéter avec la surface vasculaire, le point le plus proche sur la ligne centrale fournit un BM qui sert de succédané local à la surface globale. Ces BM sont potentiellement différents en chaque point mais il est aisé d'imposer, lors de l'avancée du micro-outil, que leur succession spatiale, aussi bien que leur mise à jour temporelle, respectent la topologie fournie par les lignes centrales.

Afin de pouvoir bénéficier du solide corpus de travaux sur la reconstruction implicite de surfaces à partir de points [BLN⁺13], nous avons besoin d'associer à chaque point de la ligne centrale des vaisseaux, un ensemble de points fournissant un échantillonnage de la surface locale. Wink et al. [WNV00] proposent de revenir à l'image pour plus de précision, après une étape initiale de suivi de vaisseaux reposant sur un modèle de section circulaire. Mais la procédure d'extraction proposée ne fournit pas un ensemble dépourvu de points aberrants, ce qui pose un sérieux problème à la reconstruction implicite, comme nous l'avons vu [MdGD⁺10]. Nous avons donc mis au point un algorithme de suivi de vaisseaux sanguins capable d'extraire leur ligne centrale, tout en associant à chaque centre un échantillonnage de la surface locale. Cet ensemble de points est construit de manière robuste grâce à l'algorithme RANSAC (*RANdom SAMpling Consensus*), ce qui nous a inspiré le nom de *RANSAC-Based Tracking* (RBT). Notre algorithme a été publié dans [21] et nous en reprenons les points essentiels en Section 4.2.

Cette étape de segmentation devait être complétée par une étape de reconstruction consistant à ajuster une surface implicite, un BM, à chaque ensemble de points local. Aucune méthode proposée pour ce faire [TBC95, BTG95] n'était assez performante pour réaliser cette reconstruction précisément et automatiquement, sans interaction manuelle, pour les milliers de cas que repré-

⁸. Stéphane Cotin a, par la suite, créé l'équipe MIMESIS à Strasbourg, dont je suis membre associé, et Christian Duriez a créé l'équipe DEFROST à Lille, que Jérémie Dequidt a rejointe.

sente un seul modèle vasculaire (un BM par point des lignes centrales). Notre algorithme, appelé *Local Implicit Modeling* (LIM), résout ce problème grâce à une formulation énergétique plus efficace du problème d'estimation et un processus de sélection plus discriminant des *blobs* (sphères implicites composant le BM). Nous précisons ces deux aspects en Section 4.3, qui reprend [20].

Enfin, nous discutons en Section 4.4 de la mise en pratique de notre modèle, et posons la question de la validation de la simulation, apportant les esquisses de réponses que nous avons dessinées dans [5, 17, 24, 26, 52].

La Section 4.5 conclut ce chapitre et suggère quelques perspectives à ce travail.

Les publications relatives à ce chapitre sont les suivantes [5, 17, 20, 21, 24, 26, 52].

Ce travail a été réalisé avec le support des projets SIMPLE⁹ (*SIMulation for PLanning the Embolization of intracranial aneurisms*), Action de Recherche Coopérative (ARC) Inria que j'ai coordonnée (2006-2008) ; SOFA, Action d'Envergure Nationale (AEN) Inria, dirigée par Stéphane Cotin, et IDEaS (*Image-Driven Simulation*), projet ANR jeune chercheur piloté par Jérémie Dequidt (2012-2016). Son cadre a été celui de la thèse d'Ahmed Yureidini [Yur14] (2010-2014), que j'ai co-encadrée, sous la direction de Stéphane Cotin, et du stage de 3^e année d'ingénieur de Pierre Glanc [Gla09] que j'ai encadré. L'essentiel de ce chapitre est repris de [5] et de nombreux détails supplémentaires sont disponibles dans la thèse d'Ahmed Yureidini [Yur14].

4.2 Segmentation vasculaire : l'algorithme RBT

4.2.1 Description générale de l'algorithme

Nous proposons de modéliser la forme locale d'un vaisseau sanguin comme un cylindre fini $\Gamma = (C, r, \vec{d}, h)$ où C est le centre, r le rayon, \vec{d} , le vecteur directeur de l'axe, et h la hauteur. Notre algorithme alterne entre des étapes de prédiction et d'estimation. L'état Γ_t^* est prédit de l'état Γ_{t-1} par :

$$C_t^* = C_{t-1} + s \frac{h_{t-1}}{2} \vec{d}_{t-1} \quad (4.1)$$

$$(r_t^*, \vec{d}_t^*, h_t^*) = (r_{t-1}, \vec{d}_{t-1}, h_{t-1}) \quad (4.2)$$

Le cylindre est donc prédit identique au précédent, si ce n'est son centre qui avance dans la direction de l'axe du cylindre d'un pas proportionnel à sa demi-hauteur par un facteur s . Nous avons fixé $s = 0.5$ afin que les cylindres successifs se superposent.

La phase d'estimation est plus complexe. Nous commençons par extraire un ensemble de points candidats à la surface du vaisseau sanguin dans un voisinage du cylindre prédit. Le meilleur cylindre $\Gamma_t = (C_t, r_t, \vec{d}_t, h_t)$ est ensuite ajusté à la forme locale par RANSAC. Nous proposons d'ajouter un sous-ensemble \mathcal{P}_t des points en consensus avec Γ_t à la description locale de la forme du vaisseau. Ces différentes étapes sont illustrées sur la Figure 4.2. Nous les détaillons dans ce qui suit.

4.2.2 Extraction des points candidats

Afin d'extraire les points candidats, N_r rayons sont lancés depuis la position courante C_t^* estimée pour le centre du cylindre, sur une distance de $3r_t^*$, et le point de gradient directionnel minimal est gardé pour chaque rayon.

Le minimum de gradient est justifié par le fait que dans les données 3DRA, et en général dans les images angiographiques, les vaisseaux sanguins sont des objets de forte intensité sur un fond sombre. Employer un minimum au lieu d'un seuil [WNV00] expose à plus de faux positifs, mais permet d'éviter un paramètre, la valeur du seuil, délicat à fixer. De plus, le gradient directionnel permet d'éviter certaines inhomogénéités dans le signal vasculaire.

La longueur de $3r_t^*$ a été fixée d'une part supérieure à la largeur du vaisseau ($\approx 2r_t^*$) afin de gérer le cas où C_t^* est proche de sa surface (ce qui peut arriver par exemple dans les sections

9. <http://simple.loria.fr>

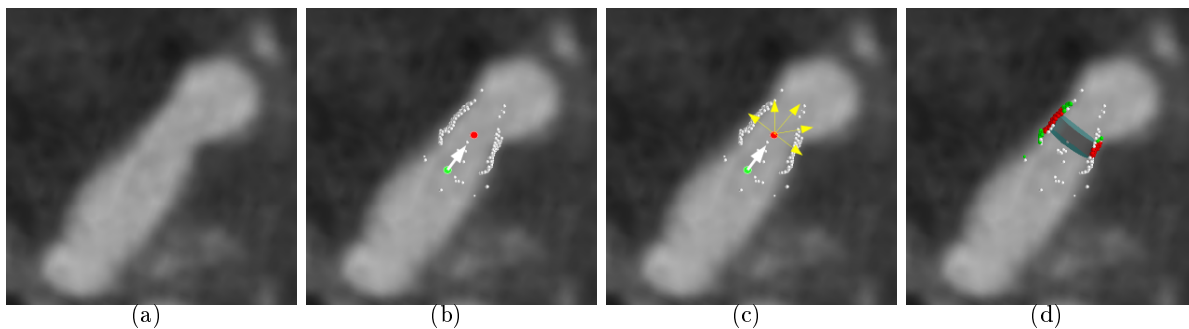


FIGURE 4.2 – Illustration des différentes étapes de RBT. (a) Coupe dans la 3DRA selon le vaisseau d'intérêt. (b) Prédiction de C_t^* (point rouge) à partir de C_{t-1} (point vert) et \vec{d}_{t-1} (flèche blanche). Des points candidats (points blancs) sont extraits à la surface du vaisseau par lancer de rayon. (c) N_d directions sont testées pour l'axe directeur du cylindre (flèches jaunes). (d) Le meilleur cylindre (C_t, r_t, \vec{d}_t) est trouvé, ainsi que son ensemble de consensus (points verts). La hauteur h_t du cylindre est estimée à partir d'un sous-ensemble de l'ensemble de consensus que nous nommons \mathcal{P}_t (points rouges).

fortement courbées), et d'autre part suffisamment grande pour gérer une variation brusque de la section vasculaire (cas pathologiques, mais aussi bifurcations).

Les directions de ces rayons sont générées en échantillonnant régulièrement la sphère gaussienne, par subdivision récursive des arêtes d'un icosaèdre initial, afin d'éviter la formation d'amas de points aux abords des pôles. Ceci arrive notamment avec un échantillonnage basé sur les coordonnées sphériques, par exemple utilisé dans la méthode des images en projection sphérique bornée (*Bounded Spherical Projection Images* [WC07]).

Enfin, différentes options ont été testées pour l'interpolation du signal nécessaire à cette extraction. Ces expériences sur des cas synthétiques ont clairement mis en avant l'importance d'une interpolation adaptée à des vaisseaux de petite taille, de diamètre équivalent à quelques voxels, et donc potentiellement sévèrement sous-échantillonnés dans les images. Le meilleur choix parmi ceux que nous avons testés consiste à trouver le point de gradient directionnel minimal en échantillonnant d'abord les données le long du rayon par approximation par B-spline cubique, puis en calculant le gradient 1D par différence centrale.

La Figure 4.2(b) illustre l'étape de prédiction (C_t^* est le point rouge), suivie de l'extraction de ces points candidats (en blanc).

4.2.3 Ajustement de cylindre par RANSAC

Le cœur de RBT consiste à ajuster un cylindre à cet ensemble de points candidats situés a priori sur la surface locale du vaisseau sanguin. Nous ne pouvons faire l'hypothèse qu'aucun de ces points n'est aberrant, certaines structures proches pouvant perturber la segmentation (par exemple deux vaisseaux qui bifurquent ou qui se touchent, problème connu comme celui du *Kissing Vessel*, voir des exemples sur la Figure 4.3). L'algorithme RANSAC (*RANdom SAMpling Consensus*) [FB81] est une méthode d'estimation robuste des paramètres d'un modèle, à partir de données, dont une certaine proportion peut être aberrante. Elle a été appliquée à l'estimation d'un cylindre à partir de points [CG01] et repose sur l'existence d'une procédure de calcul des paramètres du cylindre à partir d'un ensemble minimal de points. Cette procédure dépend de la paramétrisation choisie pour le cylindre [BF06]. Or le nombre minimal de tirages requis par RANSAC croît exponentiellement avec la taille de l'ensemble minimal de points. Nous avons donc choisi un ensemble minimal de 3 points. Ce choix revient à paramétrer le cylindre par son centre et son rayon, et suppose que la direction de son axe est connue. La procédure élémentaire de calcul des paramètres revient alors à faire passer un cercle par les points projetés dans un plan orthogonal à la direction du cylindre.

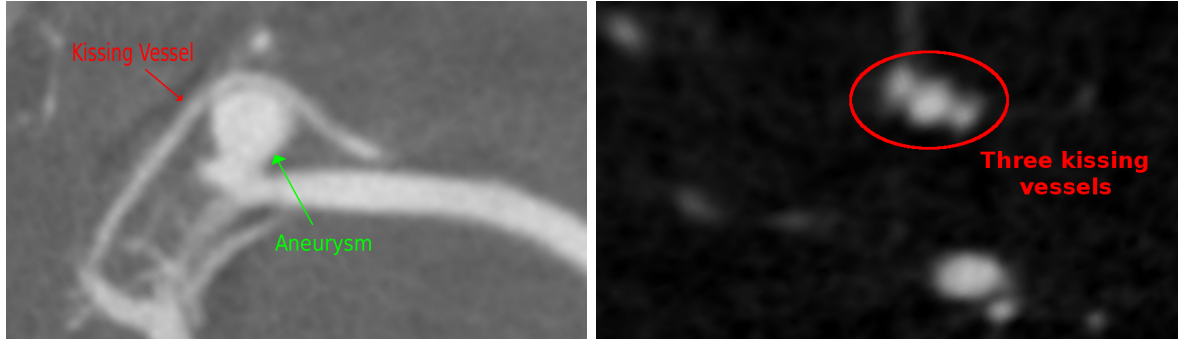


FIGURE 4.3 – Deux plans de coupe dans des 3DRA illustrant le problème *Kissing Vessel*. (à gauche) L'anévrisme a grossi au point de toucher et pousser sur un vaisseau proche. (à droite) Trois petits vaisseaux sanguins suivent la même circonvolution du cerveau et se touchent.

La direction du cylindre peut être estimée a priori à partir des normales aux points échantillonnés [CG01]. Dans notre cas, ces normales seraient calculées à partir des gradients observés dans l'image, et qui sont par conséquent bruités. Nous avons donc plutôt opté pour une stratégie similaire à [FHKP10] : N_d directions sont testées pour l'axe du cylindre sur une demi-sphère centrée sur \vec{d}_t^* , en commençant par \vec{d}_t^* puis en s'en éloignant progressivement avec un angle croissant, ce qui favorise les faibles courbures de vaisseau (voir la Figure 4.2(c)). Le meilleur cylindre est trouvé par RANSAC successivement pour chaque direction et le premier cylindre valide est conservé. Un cylindre est considéré comme valide s'il présente un taux d'inlier minimal (70% dans nos expériences) et si son rayon est dans l'intervalle $[r_t^*/2, 3r_t^*/2]$. Cette dernière condition se retrouve dans [WC07] et permet d'éviter les cylindres dégénérés (de rayon infini), tout en autorisant une variation de rayon suffisante pour gérer une grande variété de taille de vaisseaux sanguins. Dans cette même optique, on ne peut pas employer un seuil d'acceptation fixe pour RANSAC. Ce seuil permet de déterminer si un point est un inlier en se basant sur sa distance au cylindre courant. Il paraît logique d'avoir un seuil plus permissif sur les grosses artères que sur les petits vaisseaux. Nous avons donc opté pour un seuil dépendant de l'estimation courante du rayon ($0.1r_t^*$). À ce stade, un cylindre candidat est décrit par un centre, un rayon, et un ensemble de points en consensus (ensemble de points verts sur la Figure 4.2(d)).

Les paramètres du cylindre, son rayon r_t , son centre C_t mais aussi sa direction \vec{d}_t , sont alors affinés en minimisant la distance quadratique moyenne à son ensemble de consensus, et l'ensemble de consensus est mis à jour. Les derniers réglages concernent la hauteur h_t du cylindre et la position de son centre C_t le long de son axe. Pour ce faire, la hauteur est déterminée comme permettant d'englober 75% des points de l'ensemble de consensus et le centre est placé à mi-hauteur du cylindre. Cet ensemble de 75% de points englobés est appelé \mathcal{P}_t (ensemble de points rouges sur la Figure 4.2(d)).

Ainsi, la procédure de suivi de vaisseau fournit à chaque étape un nœud créé par ajustement local d'un cylindre. Ce nœud contient les informations $\{C_t, r_t, \vec{d}_t, h_t, \mathcal{P}_t\}$. Il est à noter que les points de \mathcal{P}_t sont approximativement situés sur un cylindre, mais à $0.1r_t$ près, ce qui permet de capturer certaines variations dans le profil du vaisseau sanguin (voir Figure 4.4)

4.2.4 Construction de l'arbre

En pratique, l'algorithme ci-dessus permet de suivre une branche de l'arbre vasculaire. Il s'est avéré très robuste à la présence de structures denses proches (un anévrisme sur la Figure 4.5(a) ou d'autres vaisseaux sanguins sur la Figure 4.5(b)), ou encore à celle de bifurcations qui n'empêchent pas le suivi de se poursuivre (voir Figure 4.5(c)), mais également à une mauvaise initialisation des paramètres (ceux du premier cylindre). Nous avons donc pu mettre en place la procédure suivante pour construire un ensemble de vaisseaux sous forme d'arbre (voir l'illustration sur la Figure 4.6). Pour chaque branche, il est demandé de placer manuellement un

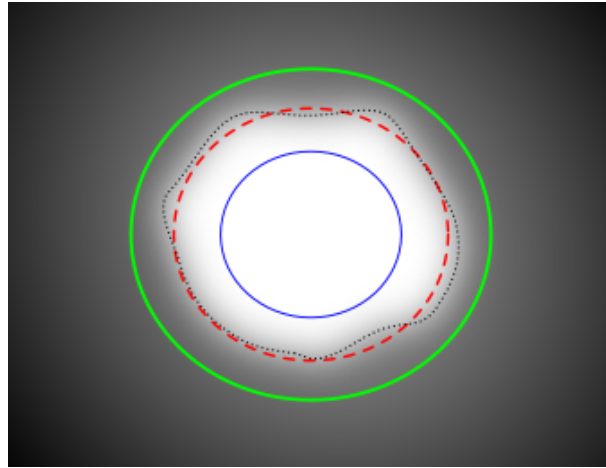


FIGURE 4.4 – Bien qu’un modèle de cylindre est ajusté, RANSAC permet d’accepter un profil de vaisseau qui s’éloigne un peu de cette forme grâce au seuil d’acceptation. Dans cette coupe de 3DRA orthogonale à un vaisseau sanguin, nous avons indiqué le cylindre trouvé de rayon r_t par le cercle rouge en pointillés. Les cercles bleu et vert représentent respectivement les limites dans lesquelles les points extraits sont considérés comme étant à la surface du vaisseau (à une distance de $0.1r_t$ du cercle rouge). Les points peuvent donc se situer sur un profil non exactement cylindrique (ligne noire en pointillés).

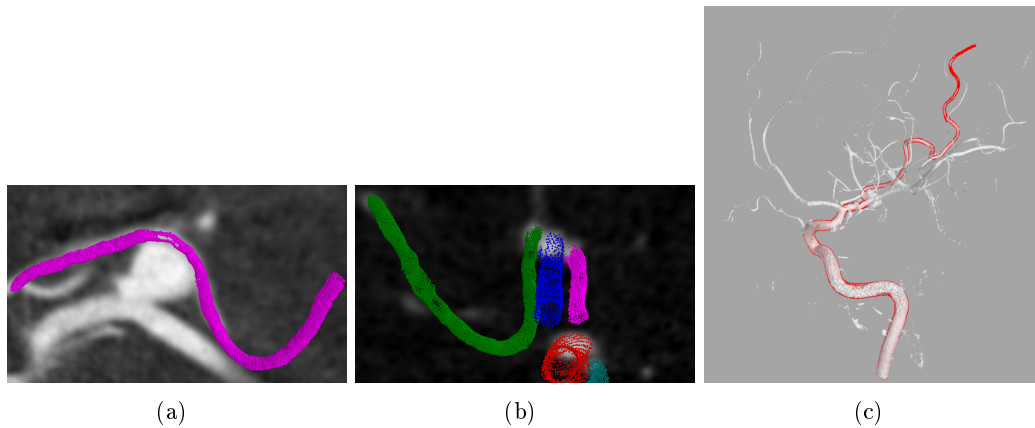


FIGURE 4.5 – Illustration de la robustesse de RBT à la présence de structures denses ((a) et (b) correspondent à la Figure 4.3) : (a) anévrisme ; (b) autres vaisseaux ; (c) mais aussi aux bifurcations qui ne stoppent pas le suivi, celui se poursuivant dans une sous-branche : dans ces exemples, le suivi est initialisé à l’origine inférieure de la carotide et se poursuit dans l’artère sylvienne au niveau du polygone de Willis, puis plus loin dans des sous-branches de la sylvienne.

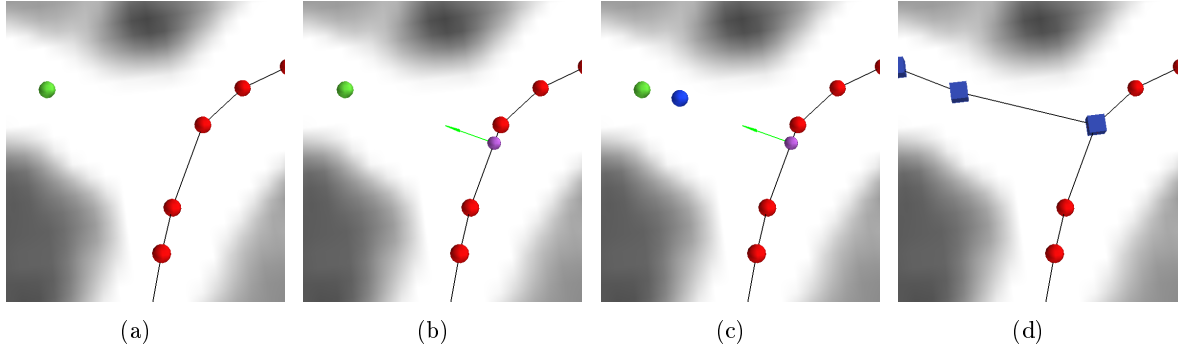


FIGURE 4.6 – Suivi d’une nouvelle branche. (a) Vaisseau parent avec sa ligne centrale (en rouge) et le point C_0^* (en vert) placé manuellement à la base du nouveau vaisseau. (b) C_0^* est projeté sur la ligne parente en C'_0 (point violet) et la direction initiale \vec{d}_0 est définie par ces deux points (flèche verte). (c) Le point initial est ajusté en C_0 (point bleu) le long de \vec{d}_0 . (d) La nouvelle branche (en bleu) est suivie et ajoutée à l’arbre courant.

point C_0^* à la base de cette branche. Si l’arbre est vide, soit un rayon et une direction initiaux sont également demandés, soit, comme en général le premier vaisseau suivi est une des quatre artères qui alimentent le cerveau (une des deux carotides ou des deux vertébrales), le rayon est fixé très approximativement suivant les données anatomiques connues (2 mm), et la direction suit l’axe vertical. Si en revanche l’arbre contient déjà des branches, on trouve la branche la plus proche de C_0^* , qui sera le vaisseau parent. Le rayon est initialisé à r_0 , valeur estimée au point C'_0 le plus proche sur le vaisseau parent. La direction \vec{d}_0 suit $C_0^* - C'_0$ après normalisation. Enfin, le point initial choisi pour débiter le suivi de la nouvelle branche est $C_0 = C'_0 + 1.25r_0\vec{d}_0$.

4.2.5 Validation

Nous avons validé RBT en comparant ses résultats sur des données patients, à ceux fournis par l’algorithme MHT (*Multiple Hypothesis Tracking*) proposé par [FHKP10] et dont les performances ont été reconnues les meilleures parmi 13 algorithmes [SMvW⁺09]. MHT est implémenté dans le module *TubularTracking* du logiciel *MevisLab*¹⁰. Nous nous sommes cependant rapidement aperçus que le mode multi-branches (permettant d’extraire l’arbre vasculaire en une seule passe) échouait sur nos données. Nous avons alors dû nous résoudre à employer le mode mono-branche, ce qui en rend l’utilisation assez semblable à RBT. De plus, l’examen visuel des résultats a rapidement mis en évidence une qualité de suivi moyenne pour MHT. Nous avons donc fait une première constatation de la difficulté particulière du suivi de vaisseaux sanguins intracrâniens par rapport aux vaisseaux coronaires considérés dans [SMvW⁺09]. Ensuite, il nous a fallu développer une approche de validation spécifique afin d’établir la meilleure performance de RBT comparé à MHT.

Il est à noter qu’un jeu de paramètres constant a été utilisé pour tous les cas traités, que ce soit pour MHT ou RBT. Le jeu de paramètres pour chaque algorithme a été établi de manière à produire un maximum de suivis corrects (maximum de suivis étiquetés SS ou AP, voir ci-dessous) sur un ensemble de 10 artères choisies aléatoirement. Les détails sont donnés dans [5].

Données cliniques

Nous avons formé notre base de données avec 10 patients accueillis au département de neuroradiologie interventionnelle thérapeutique et diagnostique du CHU de Nancy (France), pour recevoir un traitement endovasculaire pour soigner un anévrisme intracrânien. Une 3DRA de la carotide interne (droite ou gauche) a été acquise pour chaque patient dans la salle vasculaire

10. <http://www.mevislab.de>

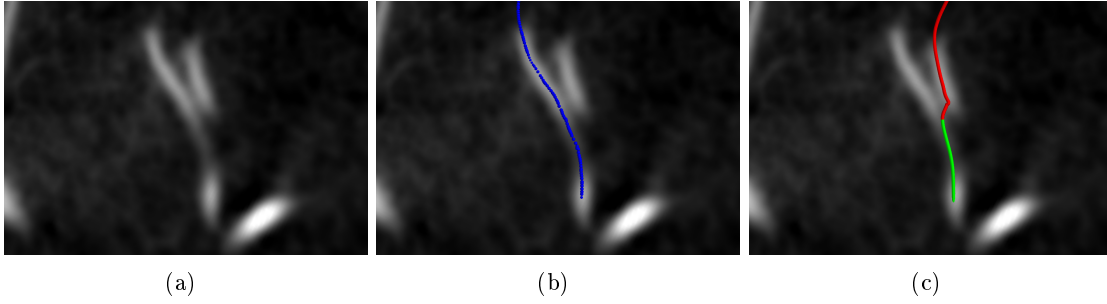


FIGURE 4.7 – Exemple de cas où RBT produit un suivi correct, mais MHT fuite dans un vaisseau sanguin proche : (a) 3DRA en coupe montrant le vaisseau suivi (gauche) et un vaisseau qui vient en contact (droite) ; (b) résultat de RBT produisant un suivi correct ; (c) résultat de MHT qui suit le mauvais vaisseau après le point de contact. Ce résultat est étiqueté EN : seule la partie verte est conservée après édition manuelle.

biplan à panneaux plats (Innova 4100, GE Healthcare). Chaque 3DRA se présente comme un cube de données de 512^3 voxels, avec une résolution comprise entre 0.18 et 0.22 mm selon le champ de vue choisi pour l’acquisition. Le diamètre des vaisseaux sanguins considérés va de 4 mm (artère carotide) à moins d’un millimètre.

Les deux algorithmes comparés (MHT et RBT) sont capables de suivre un vaisseau au-delà des bifurcations. Cependant, le choix de la branche fille n’est pas forcément le même, ce qui rend la comparaison impossible. Nous avons donc défini un *vaisseau* comme une partie d’artère entre deux bifurcations. Chaque vaisseau a été déterminé manuellement par un point initial C_0 et un point final C_f . La précision sur ces points n’est pas cruciale : C_0 est utilisé pour initialiser le suivi, et l’important est que les deux méthodes aient la même initialisation ; et les lignes centrales extraites par suivi sont coupées au point le plus proche de C_f (si à une distance inférieure à 2 fois le rayon estimé r_t , sinon, le suivi a échoué et l’édition est manuelle, voir plus bas). Dès lors, il a simplement été vérifié que deux experts positionnaient C_f au niveau de la même bifurcation. En cas de désaccord, le point C_f désigné le plus proximal¹¹ était conservé, l’hypothèse étant que dans le pire cas, un expert au moins détectait la bonne bifurcation.

745 vaisseaux ont ainsi été définis à partir des 10 volumes 3DRA, avec entre 57 à 96 vaisseaux par patient. Les vaisseaux correspondant aux artères carotides (jusqu’au polygone de Willis) ont été retirés de l’évaluation car trop larges pour être segmentés par MHT (le suivi fonctionne sans problème pour RBT).

Mesures de performance qualitatives

Le résultat du suivi sur chaque vaisseau sanguin a été étiqueté indépendamment par deux experts, puis un consensus a été établi pour les quelques cas de désaccord. Les étiquettes sont :

- Succès du Suivi (SS) : C_f est atteint et aucune édition de la ligne centrale n’est nécessaire ;
- Arrêt Prématuro (AP) : C_f n’est pas atteint mais la ligne centrale reste visuellement dans le vaisseau sur toute sa longueur ;
- Suivi Erroné (SE) : aucune ligne centrale n’est produite ou bien un vaisseau proche ou parent est suivi dès le départ ;
- Édition Nécessaire (EN) : les lignes centrales restantes sont correctes au départ, puis une erreur se produit : le suivi accroche une structure dense proche (voir la Figure 4.7), ou se poursuit dans le fond de l’image (bruité), ou effectue un demi-tour et remonte le vaisseau.

Ces cas ont été édités à la main, la ligne centrale étant coupée à la première erreur (voir Figure 4.7). Une analyse plus fine de ces cas est présentée dans [5].

L’évaluation établit les pourcentages de cas pour chaque étiquette. Ces scores sont reportés dans le Tableau 4.1. Ils démontrent un taux de succès bien supérieur pour RBT (82.5% de SS comparés

11. Le plus proche de C_0 le long de la ligne centrale

	SS (%)	AP (%)	SE (%)	EN (%)	SS+AP (%)	LS (mm)	SS + AP + EN (%)	LSE (mm)
MHT	59.87	0.27	9.66	30.2	60.14	5959	90.34	8407.8
RBT	82.55	12.89	4.16	0.40	95.44	10909.5	95.84	11637.7

TABLE 4.1 – Distribution des étiquettes des vaisseaux, en pourcentage, après suivi par MHT et RBT. SS=Succès du Suivi, AP=Arrêt Prématuré, SE=Suivi Erroné, EN=Édition Nécessaire. Les vaisseaux étiquetés SS ou AP (SS+AP) ne comportent pas de partie erronée et ont été utilisés pour calculer la Longueur Suivie (LS, en mm). Les vaisseaux étiquetés EN ont été édités manuellement pour ne préserver que la partie proximale correcte (voir Figure 4.7(c)) et ont été utilisés, avec les suivis SS et AP (SS+AP+EN), pour calculer la Longueur Suivie après Édition (LSE, en mm). (Résultats sur 745 vaisseaux, voir le texte pour plus de détails)

au 50.9 pour MHT) et une tendance à stopper un suivi (12.9% de AP) plutôt que de poursuivre en suivant une ligne centrale erronée. C’est le contraire pour MHT dont les résultats nécessitent fréquemment une édition (30% de SE).

Mesures quantitatives

La longueur cumulée des lignes centrales extraites a été calculée pour chaque algorithme. Cette mesure vient compléter les pourcentages mesurés ci-dessus, en ce sens qu’elle est aussi un indicateur de la capacité d’un algorithme à suivre des vaisseaux dans différentes conditions et configurations. Cette Longueur de Suivi (LS) est reportée dans le Tableau 4.1 où elle est calculée à partir des seuls suivis qui ont réussi (étiquetés SS ou AP). La LS est presque deux fois plus élevée pour RBT, mais la comparaison n’est pas totalement juste puisque seuls 60% des suivis sont comptabilisés pour MHT contre 95.4 pour RBT (SS+AP). Nous avons donc calculé une nouvelle longueur de suivi qui intègre cette fois-ci les lignes centrales éditées (SS+AP+EN). Nous montons ainsi à des taux plus similaires de 90.3% pour MHT et 95% pour RBT de suivis comptabilisés. Malgré cet ajustement, la mesure, nommée Longueur de Suivi après Édition (LSE), reste plus de 38% plus grande pour RBT comparée à celle associée à MHT.

Les critères précédents ne disent rien sur la précision du suivi obtenu par RBT. En supposant que MHT, dans les cas de succès, produit un résultat qui peut servir de vérité terrain, nous avons donc mesuré l’erreur induite par RBT en calculant le critère classique de la distance symétrique moyenne. Ce critère ne peut a priori être calculé que sur les lignes centrales éditées. Cette manière de faire a deux inconvénients. D’une part elle peut être biaisée par la qualité de l’édition manuelle (à partir de quand la divergence de suivi au vaisseau est-elle assez grande pour être erronée?). D’autre part, il est apparu que dans un certain nombre de cas MHT récupérait d’une erreur locale et produisait un suivi correct en aval de cette erreur, partie en aval que l’édition supprime. La Figure 4.8 donne un exemple de cas pour le suivi d’une artère péricalléuse. Nous avons donc utilisé la Distance Symétrique Moyenne Modifiée (DSMM) :

$$DSMM = \frac{1}{N_w + M_w} \left(\sum_i^N d_w(P_i, \mathcal{Q}) + \sum_j^M d_w(Q_j, \mathcal{P}) \right) \quad (4.3)$$

où $\mathcal{P} = \{P_i\}_{i \in [1, N]}$ est la ligne centrale produite par MHT, et $\mathcal{Q} = \{Q_j\}_{j \in [1, M]}$ celle produite par RBT. d_w est la distance euclidienne entre un point et une courbe, seuillée afin de renvoyer 0 dès lors qu’elle dépasse w . N_w et M_w sont le nombre de points pris en compte sous le seuil dans respectivement \mathcal{P} et \mathcal{Q} . Le principe de la DSMM est ainsi de sélectionner automatiquement les portions communes aux deux lignes centrales et renvoie la distance symétrique moyenne entre elles. La DSMM est majorée par w et l’on pourrait bien évidemment obtenir un résultat en apparence excellent en choisissant une valeur petite pour w . Afin de montrer l’absence de biais dans nos résultats, nous avons calculé la DSMM pour deux valeurs de w : une première, fixe, de 3 voxels (ce qui correspond à un vaisseau de diamètre entre 1 et 1.5 mm) et une deuxième, variable,

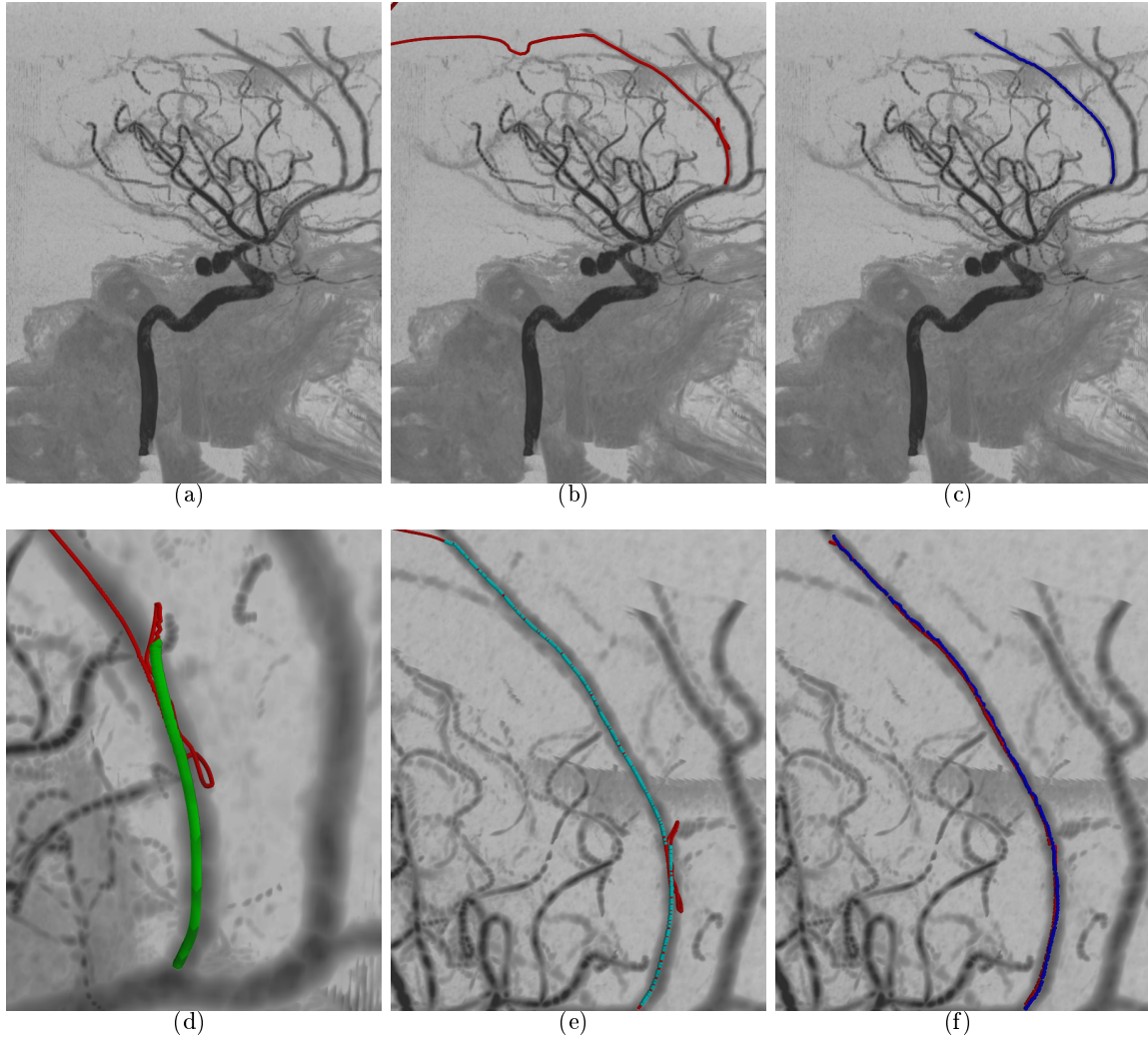


FIGURE 4.8 – Intérêt de la Distance Symétrique Moyenne Modifiée (DSMM) sur un exemple. (a) Vue MIP de la 3DRA. (b) Résultat du suivi de l'artère péricalleuse par MHT : le suivi ne s'arrête pas au bon endroit, et on peut remarquer une erreur (sous forme de nœud) dans la partie proximale. (c) Résultat du suivi par RBT : la ligne centrale est correcte. (d) Zoom de (b) sur l'erreur proximale en forme de nœud : la ligne centrale initiale obtenue par MHT (en rouge) est correcte dans sa partie proximale. Le cas est donc étiqueté EN et on ne conserve que la partie verte pour le calcul de la LSE, négligeant alors la partie, pourtant correcte, en aval de l'erreur. (e) Les points sélectionnés automatiquement pour le calcul de la DSMM sont montrés en cyan. On remarque que seules les parties visuellement correctes sont conservées, avant et après l'erreur, et les points erronés sont bien exclus ($w = \bar{r}$) (f) Points utilisés pour calculer la DSMM : MHT en rouge, RBT en bleu.

de \bar{r} , rayon moyen estimé sur le vaisseau suivi. De cette manière, des résultats similaires, et bien inférieurs à w doivent prouver la capacité de la DSMM à sélectionner de manière cohérente les portions communes aux lignes centrales, et donc produire un résultat exploitable.

Les résultats obtenus vont dans ce sens puisque la DSMS a été mesurée à 0.18 mm sans différence (au centième de mm près) entre les deux valeurs de w . La mesure est systématiquement légèrement inférieure à la taille voxel (min=0.16 mm et max=0.21mm pour la DSMM avec une taille voxel entre 0.18 et 0.22 mm). RBT produit donc des lignes centrales avec une précision équivalente à MHT.

4.3 Reconstruction vasculaire : l'algorithme LIM

RBT calcule une segmentation de l'arbre vasculaire sous la forme d'un arbre qui reproduit la topologie des artères et où chaque nœud contient d'une part, des informations sur la ligne centrale en termes de position (centre C), de direction (\vec{d}) et en quelque sorte de courbure à travers la hauteur (h) du cylindre ajusté localement ; et d'autre part des informations sur la forme locale du vaisseau par une estimation de son rayon (r) et un échantillonnage local de sa surface (\mathcal{P}). Ces informations de forme, bien qu'elles soient utilisées par plusieurs travaux antérieurs dans la littérature (voir la fin de la Section 4.1.1), ne suffisent pas à représenter la surface vasculaire de manière assez précise et complète pour mener une simulation interactive, et surtout réaliste. Nous procédons donc à une étape supplémentaire de reconstruction de la surface vasculaire afin de générer les informations nécessaires à la gestion des collisions localement à chaque nœud. Ces informations prennent la forme d'une surface implicite locale, raison pour laquelle nous avons nommé notre algorithme Modélisation Implicite Locale ou *LIM* pour *Local Implicit Modeling*. Les différentes étapes de LIM sont illustrées sur la Figure 4.9, et s'appliquent à chaque nœud (par exemple un nœud de bifurcation dans la Figure 4.9(b)) :

- sélection des voisins topologiques (Figure 4.9(c)) ;
- concaténation des ensembles de points associés à ces nœuds (Figure 4.9(d)) : ainsi les ensembles successifs sont recouvrants afin de favoriser la continuité des jonctions entre surfaces implicites locales adjacentes ;
- initialisation de la surface implicite (Figure 4.9(e)) : notre méthode s'accommode d'une initialisation grossière ;
- optimisation de la surface implicite pour s'adapter aux points (Figure 4.9(f)) : LIM est robuste aux trous, mais suppose qu'aucun point n'est aberrant, cette propriété étant assurée suite au filtrage par RANSAC dans RBT.

Nous reprenons ici les éléments essentiels concernant le modèle implicite choisi, l'expression de l'objectif d'ajustement de la surface aux points sous la forme d'une énergie à minimiser, la procédure d'optimisation et la validation de LIM. À nouveau, des précisions supplémentaires sont disponibles dans [Yur14, 5].

4.3.1 Modèle implicite

Les surfaces implicites s'appuyant sur des squelettes sont particulièrement bien adaptées à la modélisation de formes quelconques et lisses, sous une expression à la fois compacte et versatile [BTG95]. De plus, et en cohérence avec notre objectif de surface implicite locale, elles possèdent une bonne propriété de localité [DvOG04]. Dans le cadre du stage de Master de Pierre Glanc [Gla09], nous nous sommes intéressés aux surfaces basées sur un ensemble de points pour squelette, et ces premiers résultats nous ont incités à poursuivre nos investigations dans ce sens pendant la thèse d'Ahmed Yureidini [Yur14].

Une surface implicite générée par un ensemble de points est formée par le lieu des points annulant une fonction f , définie par la somme finie de N_b sphères implicites [Bli82] :

$$f(P) = \sum_{j=1}^{N_b} \alpha_j \phi \left(\frac{|P - C_j|}{\rho_j} \right) - T \quad (4.4)$$

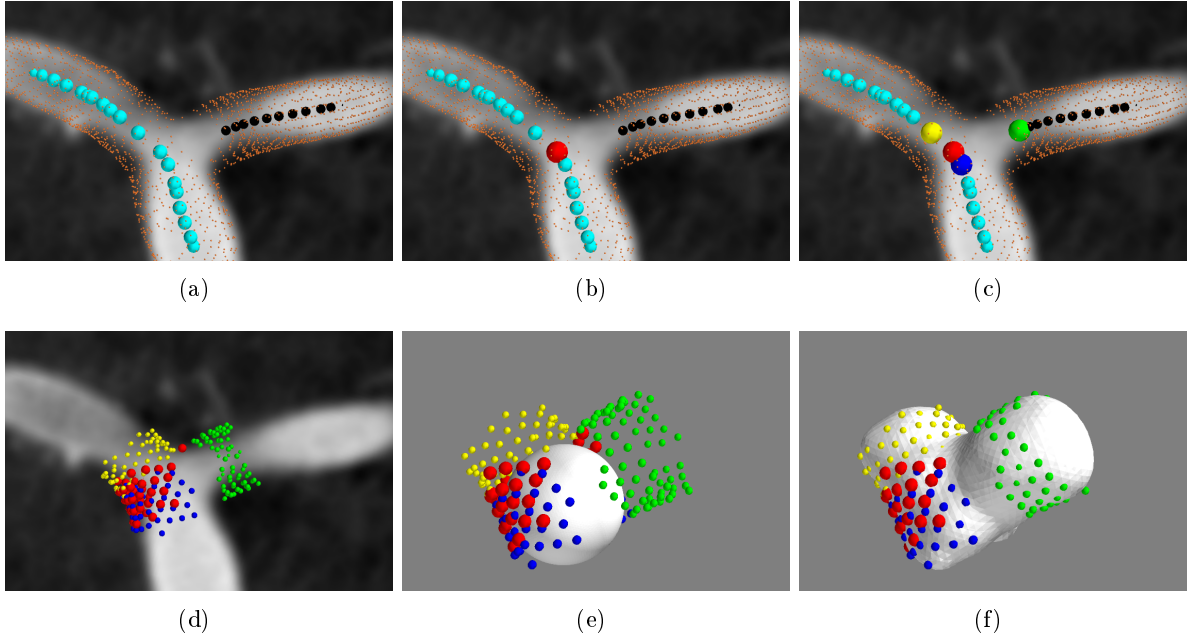


FIGURE 4.9 – Principales étapes de l’algorithme LIM. (a) Données en entrée (zoom sur une bifurcation) : centre des nœuds de l’arbre RBT (deux branches : points bleus et noirs), et points échantillonnés à la surface vasculaire (points orange, les ensembles de points pour tous les nœuds sont fusionnés sur cette vue). (b) Le nœud de bifurcation (centre rouge) est choisi pour la modélisation. (c) Les voisins topologiques sont sélectionnés (leurs centres sont indiqués en jaune, bleu et vert). (d) Un ensemble de points 3D est formé par concaténation des ensembles locaux aux nœuds sélectionnés. L’objectif de LIM est d’estimer une surface implicite qui passe au mieux par ces points. (e) On ne conserve que les points et la surface implicite est initialisée (un BM formé de deux sphères implicites proches). (f) Résultat final de LIM qui modélise avec succès la surface locale à la bifurcation (après 30 subdivisions). L’intérêt de la concaténation des ensembles de points locaux réside dans une meilleure capture de la forme locale du vaisseau, et une meilleure continuité de jonction des BM adjacents.

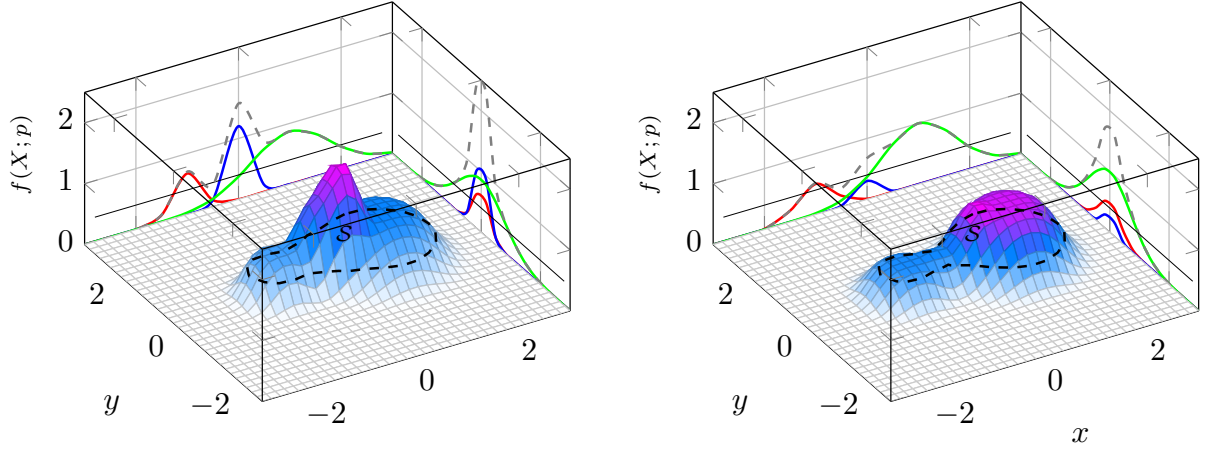


FIGURE 4.10 – Illustration de la redondance des paramètres des BM. Deux champs scalaires f différents, définis sur \mathbb{R}^2 , mais tous deux générés à partir de trois gaussiennes, produisent pour le même seuil $T = 0.4$ deux courbes implicites identiques : à gauche, le paramétrage est $\alpha = \{0.6715, 1.2006, 0.8681\}$ et $\rho = \{0.3722, 0.3258, 1.0863\}$; et à droite, nous avons imposé $\alpha = \rho = \{0.3722, 0.3258, 1.0863\}$. On peut remarquer que dans ce dernier cas, la fonction f est une meilleure indicatrice de la fonction distance à la forme.

où $P \in \mathbb{R}^3$, $T \in \mathbb{R}^+$ est le seuil pour ajuster l'isosurface, et les points $\{C_j\}$ définissent le squelette. Chaque sphère implicite d'indice $\#j$ est définie par une fonction à symétrie sphérique positive ϕ appelée *profil* ou *noyau* telle que $\phi(0) = 1$ et rapidement décroissante vers 0 en l'infini [She99]. Le rayon de la sphère est modulé à travers sa *largeur* ρ_j qui permet de changer l'échelle du profil, et sa position est définie par translation du profil en son *centre* C_j . Enfin, la sphère est pondérée par α_j , assimilable à une *hauteur* puisque c'est la valeur que prend le blob en C_j .

L'appellation de ces modèles varie en fonction du noyau utilisé [She99]. En référence au travail pionnier réalisé par Muraki [Mur91], nous utiliserons le terme *blob* pour une sphère implicite et *Blobby Model* (abrégé en BM dans la suite) pour le modèle implicite basé sur f . Notre implémentation des BM est agnostique à la fois quant à la dimension et au profil. Nous avons donc pu travailler en 2D, où l'analyse visuelle est plus aisée, et en 3D, où sont définies nos données 3DRA. Nous avons aussi pu tester divers profils (gaussien, de Cauchy, metaballs, *soft objects*, quartique [She99]) sans détecter de différence majeure sur le résultat. Toutes nos expériences ont donc été faites en utilisant la fonction de Cauchy $\phi(r) = (1 + r^2/5)^{-2}$ (à des fins de comparaison, tous nos profils ont été normalisés de manière à ce que $\phi''(1) = 0$). Ce profil, inverse d'un polynôme, est rapide à calculer et possède un support infini, ce qui simplifie la gestion des points de gradient nul.

L'expression de l'Équation 4.4 définit des surfaces implicites procédurales, très efficaces d'un point de vue calculatoire. Elle présente cependant une redondance dans ses paramètres, ce qui permet à une seule surface de correspondre à plusieurs paramétrages (Figure 4.10). Une solution classique impose pour tout j $\rho_j = 1$, ce qui mène aux fonctions à base radiale (si les centres sont fixes), mais deux arguments nous ont orientés vers le choix différent d'imposer $\alpha_j = \rho_j$. Ces arguments reposent sur l'intérêt, dans notre contexte de simulation, que f varie de façon monotone avec la distance à la surface. Ceci facilite la gestion des collisions en reliant la distance d'interpénétration à la valeur de f , et la direction de reprojection d'un point sur la surface au gradient de f .

Le premier argument considère un BM ne comportant qu'une seule sphère. La distance à la surface au centre de cette sphère est idéalement ρ , alors que la fonction renvoie α .

Le deuxième argument est plus subtil, mais a plus de poids en pratique. Un préalable à son établissement dissèque le profil d'un blob (nous posons $\rho = 1$, sans perte de généralité à ce stade). Ce profil prend la forme $\phi(r)$ avec $r = |P - C|$. Le processus de collision va requérir le gradient

de f par rapport à P , qui repose sur

$$\nabla_P \phi(r) = \phi'(r) \nabla_P r = \phi'(r) \frac{P - C}{|P - C|} \quad (4.5)$$

Nous faisons donc face à une singularité en $P = C$. Mais nous pouvons remarquer que tous les profils classiquement utilisés dépendent en fait de r^2 , de telle manière qu'il existe une fonction auxiliaire ψ telle que $\psi(r^2) = \phi(r)$ (revoir l'expression du noyau de Cauchy plus haut). Ceci permet d'éviter la singularité car dans ce cas, le gradient est :

$$\nabla_P \psi(r^2) = 2\psi'(r^2)(P - C) \quad (4.6)$$

L'expression de f devient alors :

$$f(P) = \sum_{j=1}^{N_b} \alpha_j \psi \left(\frac{|P - C_j|^2}{\rho_j^2} \right) - T \quad (4.7)$$

Et si, pour gérer les collisions, nous en formons le gradient, nous obtenons :

$$\nabla_P f(P) = 2 \sum_{j=1}^{N_b} \frac{\alpha_j}{\rho_j} \psi' \left(\frac{|P - C_j|^2}{\rho_j^2} \right) \frac{P - C_j}{\rho_j} \quad (4.8)$$

Poser pour tout j que $\alpha_j = \rho_j$, offre donc une expression équilibrée pour le gradient où l'influence de chaque blob est normalisée, sans impact de sa hauteur ou de sa largeur.

L'expression de f utilisée dans cette étude est ainsi :

$$f(P) = \sum_{j=1}^{N_b} \rho_j \psi \left(\frac{|P - C_j|^2}{\rho_j^2} \right) - T \quad (4.9)$$

Ce BM est paramétré par les centres $\{C_j\}$ et les largeurs $\{\rho_j\}$, soit un vecteur $\{C_j^T, \rho_j\}_{j=1\dots N_b}$ de taille $4N_b$. On notera dans la suite, pour tout j , la fonction définissant un blob :

$$f_j(P) = \rho_j \psi \left(\frac{|P - C_j|^2}{\rho_j^2} \right) \quad (4.10)$$

4.3.2 Formulation énergétique

L'ajustement d'une surface à un ensemble de N_p points $\{P_i\}$ est formulé de manière classique comme la minimisation d'une énergie [Mur91, BTG95, TBC95]. Cette énergie comporte un terme d'attache aux données et un terme de régularisation.

Nous avons choisi d'exprimer l'attache aux données par :

$$\mathcal{E}_d = \frac{1}{N_p} \sum_i f(P_i)^2 \quad (4.11)$$

Cette énergie exprime le fait que les points $\{P_i\}$ doivent se retrouver sur l'isosurface nulle de f , en minimisant leur distance algébrique à cette surface. En revanche, elle ne spécifie pas quelle doit être l'orientation de la surface implicite. Dans le cas où les données en entrée sont des points orientés, cette orientation peut être déterminée en maximisant l'alignement des normales en ces points avec le gradient de f , par minimisation d'un terme supplémentaire d'énergie [Mur91], ce qui équivaut au problème de Poisson [KBH06]. Nous avons établi expérimentalement que ce terme n'avait aucun effet dans notre cas où les points sur la ligne centrale sont à l'intérieur du vaisseau et la surface peut donc être initialisée à l'intérieur de l'espace englobé par les points en entrée [43, Yur14].

Diverses formulations ont été proposées pour le terme de régularisation. La régularisation de Tikhonov minimise l'intégrale de la norme du gradient, afin de favoriser des surfaces lisses, mais

des termes d'ordre supérieur sont nécessaires pour éviter l'effet "collier de perles" [Lem10] observé en employant des sphères implicites comme briques de base. Ces termes visent à minimiser la courbure totale de la surface. Or, une formule directe existe pour calculer la courbure moyenne d'une fonction implicite en tout point P [Gol05] :

$$\kappa(P) = \frac{\nabla f^t H_f \nabla f - |\nabla f|^2 \text{trace}(H_f)}{2|\nabla f|^3} \quad (4.12)$$

dans laquelle ∇f est le gradient de f , et H_f sa matrice hessienne, tous deux au point P . Cela nous conduit au terme de lissage, qui est lié au critère d'aire minimale :

$$\mathcal{E}_a = \frac{1}{N_p} \sum_i \kappa(P_i)^2 \quad (4.13)$$

Un autre terme de régularisation couramment employé permet d'éviter que le volume du BM ne s'annule [Mur91, TBC95]. Un point de vue particulièrement intéressant regarde le BM comme un système de particules (les blobs) dont on va éviter l'effondrement sur un seul point grâce à une force interne répulsive dépendant de la distance inter-particules [WH94]. Mais un autre cas dégénéré est celui du BM dont tous les blobs sont infiniment distants les uns des autres. Cette déliquescence du BM, en contraction comme en extension, peut se gérer grâce à des forces internes inter-particulaires qui sont attractives à grande distance et répulsives à faible distance [SBS03]. Ces forces dérivent du potentiel de van der Waals, bien connu en chimie, dont une approximation mathématique plus simple à manipuler est fournie par le potentiel de Lennard Jones. Cependant ce dernier se fonde sur l'hypothèse que les particules ont toutes le même diamètre, ce qui n'est pas notre cas, puisque chaque blob possède sa propre largeur. Nous proposons donc d'adapter ce potentiel à des particules de rayons hétérogènes et définissons l'énergie interne :

$$\mathcal{E}_c = \frac{1}{N_b(N_b - 1)} \sum_{j \neq k} \left(\frac{s\sqrt{\rho_j \rho_k}}{|C_j - C_k|} \right)^{12} - 2 \left(\frac{s\sqrt{\rho_j \rho_k}}{|C_j - C_k|} \right)^6 \quad (4.14)$$

où chaque terme de la somme atteint sa valeur minimale de 1 quand $|C_j - C_k| = s\sqrt{\rho_j \rho_k}$, autrement dit la distance entre deux centres est proportionnelle à la moyenne géométrique des largeurs. Nous avons fixé s à 2 dans toutes nos expérimentations, sur la base que deux blobs de même largeur ρ seront séparés de 2ρ à la position du minimum. \mathcal{E}_c produit une force répulsive pour les blobs plus proches que cette distance, et attractive s'ils sont plus éloignés.

La forme finale de l'énergie combine ces trois termes en les pondérant :

$$\mathcal{E} = \mathcal{E}_d + \alpha \mathcal{E}_a + \gamma \mathcal{E}_c \quad (4.15)$$

où α et γ sont des poids positifs. Ces trois termes, parmi d'autres, ont été reconnus expérimentalement à la fois comme complémentaires et effectifs dans l'ajustement de la surface implicite [43, Yur14]. Il est important de remarquer que l'énergie peut se calculer de manière algébrique, et qu'il en est de même pour ses gradients (par rapport à P , ainsi qu'aux paramètres $\{C_j\}$ et $\{\rho_j\}$), même si le calcul est trop long pour être reproduit dans une publication, sans toutefois être trop difficile, si ce n'est qu'il fait appel au calcul tensoriel pour \mathcal{E}_a . Ces calculs sont esquissés en annexe de [Yur14]. Nous employons ces expressions algébriques dans notre implémentation.

4.3.3 Optimisation

La connaissance des gradients de l'énergie permet de tirer profit d'une minimisation par descente de gradient. Cependant, le grand nombre de paramètres à optimiser, même pour des BM de taille raisonnable (40 paramètres pour un BM de 10 blobs), rend la convergence trop lente, sachant que plusieurs milliers de BM doivent être ajustés pour modéliser un arbre vasculaire. À ce problème s'ajoute celui, classique, de déterminer le nombre de paramètres, autrement dit de blobs, du BM, ce qui s'apparente à la sélection de modèles. Cependant, la modélisation de l'erreur n'est pas triviale et il reste le problème de générer les modèles candidats. Nous avons

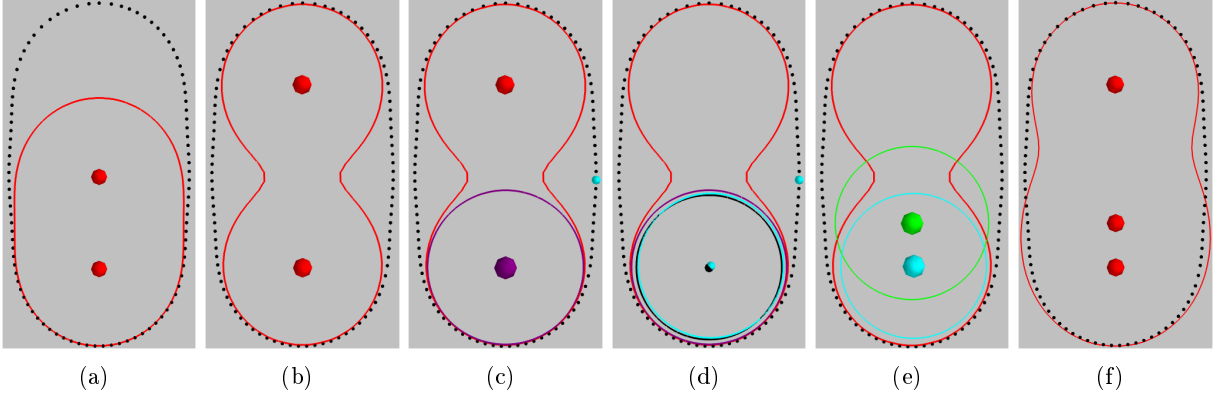


FIGURE 4.11 – Étapes élémentaires de sélection-subdivision lors de l’optimisation d’un BM dans LIM (en 2D). Les points en entrée \mathcal{P} sont en noir, le contour implicite courant en rouge et les sphères représentent les centres des blobs. (a) Initialisation du BM avec deux blobs. (b) Résultat de la première optimisation par descente de gradient sur les centres puis les rayons des blobs. (c) Détermination du point le plus éloigné du contour implicite (cyan) et sélection du blob dont il est le plus proche (violet). (d) Subdivision en deux du blob sélectionné : un blob reste au même endroit (violet) et l’autre est décalé légèrement vers le point le plus éloigné (cyan). La largeur des deux blobs est choisie telle que la somme des deux blobs conduit au même contour implicite que le blob original afin de préserver le contour courant (en rouge). (e) Optimisation du centre puis de la largeur du nouveau blob décalé (vert). (f) Nouveau contour implicite à l’issue de cette itération (rouge). L’itération suivante reprend en (c) jusqu’à convergence.

donc suivi une stratégie classique de complexification progressive du modèle : partant d’un BM minimal (un ou deux blobs), notre approche sélectionne le blob le plus pertinent, le subdivise en deux, et optimise le nouveau BM, ces opérations étant itérées jusqu’à ce qu’un objectif d’erreur maximale soit atteint. La Figure 4.11 illustre les étapes itérées lors de l’optimisation.

Cette approche est déjà présente dans [Mur91], mais la sélection du blob à subdiviser suit une approche exhaustive très gourmande en temps de calcul. Un critère de sélection basé sur la distance algébrique (valeur de \mathcal{E}_d sur un voisinage local) a été proposé dans [TBC95]. D’après nos expériences, ce critère a cependant tendance à favoriser les petits blobs, et focalise par conséquent l’optimisation d’abord sur des détails avant d’assurer une bonne forme globale pour la surface. Nous avons donc proposé d’employer plutôt un critère basé sur la distance géométrique dont une approximation a été donnée pour les surfaces implicites par [Tau91] :

$$d_T(P, f) = \frac{|f(P)|}{|\nabla f(P)|} \quad (4.16)$$

Grâce à cela, nous commençons par déterminer le point le plus éloigné de la surface implicite (point couleur cyan sur la Figure 4.11(c)) d’indice i^* tel que :

$$i^* = \arg \max_{1 \leq i \leq N_p} d_T(P_i, f) \quad (4.17)$$

Puis nous sélectionnons le blob d’indice j^* dont l’isosurface s’approche le plus de P_{i^*} (blob le plus pertinent, en violet sur la Figure 4.11(c)) :

$$j^* = \arg \min_{1 \leq j \leq N_b} d_T(P_{i^*}, f_j) \quad (4.18)$$

f_j étant donné par l’Équation 4.10. Une étape de subdivision suit où le blob sélectionné est remplacé par deux blobs de même centre que le blob original, et de largeur telle que la somme des deux nouveaux blobs produit la même surface implicite que le blob original (blobs noir et cyan sur la Figure 4.11(d)). Afin d’éviter une valeur infinie pour \mathcal{E}_c , un léger décalage est appliqué à un des deux blobs, de $\rho_{j^*}/10$ vers le point le plus proche identifié (blob cyan sur la Figure 4.11(d)).

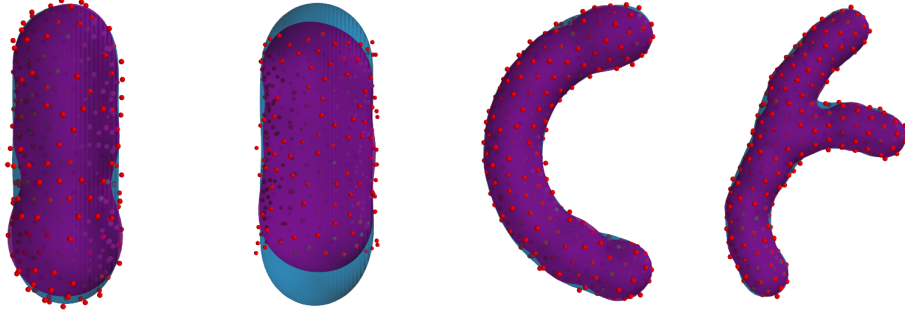


FIGURE 4.12 – Formes synthétiques utilisées pour valider LIM. La surface de référence est en bleu. Le BM issu de LIM est en violet : ce résultat correspond au pire cas obtenu pour chaque forme sur les 10 niveaux de bruit gaussien testés (entre 0.025 et 0.25 mm d'écart-type, 30 instances par niveau). Les points bruités donnés en entrée de LIM sont en rouge.

Les paramètres de ce dernier blob sont optimisés (blob vert sur la Figure 4.11(e)) avant de reprendre les étapes de sélection puis subdivision. Cette focalisation sur un seul blob est autorisée grâce aux propriétés de localité des blobs dans le BM. Les coordonnées du centre étant difficilement commensurables avec le rayon, nous avons opté pour une séparation de ces deux paramètres, alternant une minimisation de l'énergie sur les coordonnées du centre, avec une minimisation sur le rayon. Parmi les variantes proposées pour la descente de gradient dans la *GNU Scientific Library* (GSL), celle de Polak-Ribière a donné les meilleurs résultats. Cette procédure est itérée jusqu'à ce que la distance du point le plus éloigné ($d_T(P_{i^*}, f)$ dans l'Équation 4.17) soit inférieure à une valeur cible t_g . En pratique, 5 itérations suffisent en général. Un nombre maximal N_s de blobs a aussi été fixé par BM afin de limiter la complexité des BM.

Chaque BM est initialisé à partir des caractéristiques du cylindre associé au nœud dans l'arbre RBT. D'une manière similaire à l'étape de subdivision, deux blobs sont générés, centrés sur le centre du nœud, de rayon tel que leur somme conduit à une sphère implicite inscrite dans le cylindre. Puis ils sont décalés légèrement dans la direction \vec{d}_t . Chaque BM est optimisé de manière indépendante des autres, ce qui permet des calculs en parallèle, mais n'assure pas une continuité formelle entre BM voisins. Celle-ci est simplement favorisée par la superposition des ensembles de points successifs utilisés comme contraintes.

4.3.4 Validation

Comme pour RBT, une campagne de validation approfondie a été menée pour caractériser les forces et limites de LIM. Un rapport détaillé est donné dans [5] et une présentation exhaustive des résultats est disponible dans [Yur14]. Nous ne reprenons ici que les points essentiels, en mettant l'accent sur la justification des choix opérés, et en illustrons les conclusions principales que nous en avons tirées.

Référence et objectifs de l'évaluation

Comme pour RBT, les deux questions de la référence et du critère se sont posées. Nous avons donc en premier lieu effectué des tests synthétiques dont l'objectif était d'une part une première validation de LIM et d'autre part de définir un objectif réaliste pour l'évaluation sur données patients. Au niveau du paramétrage de LIM, les poids α et γ des énergies ont été fixés sur quelques exemples afin de donner un même ordre de grandeur aux trois termes énergétiques ($\alpha = 10^{-3}$ et $\gamma = 10^{-5}$), et le seuil T pour l'isosurface a été défini à la moitié de la taille voxel (0.1). Les valeurs cibles pour l'optimisation, t_g et N_s , ont été étudiées expérimentalement.

Les quatre formes synthétiques utilisées sont illustrées sur la Figure 4.12. Les points utilisés en entrée de LIM ont été échantillonnés sur la surface de référence [WH94], puis bruités sur 10

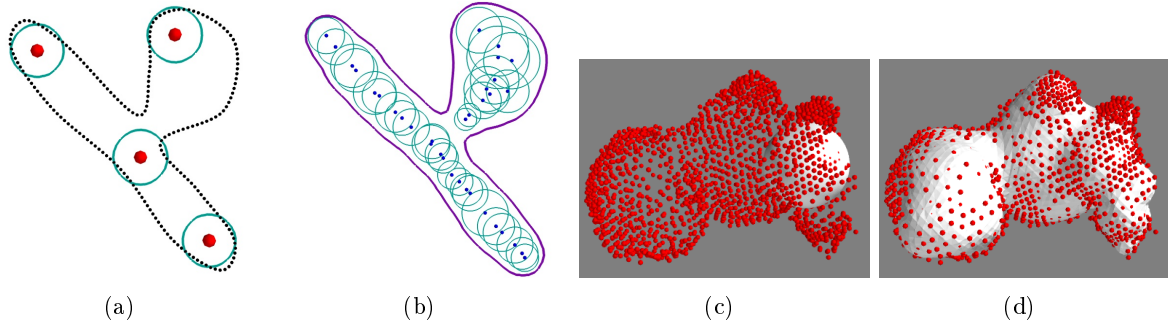


FIGURE 4.13 – Illustration des performances de LIM sur des cas complexes en 2D ((a) et (b)) et 3D ((c) et (d)). (a) Les points en entrée sont en noir et le BM est initialisé avec 4 blobs (cercles verts avec centres en rouge). (b) Résultat final, le contour implicite (violet) s’ajuste précisément aux points, et les blobs du BM final (cercles verts) sont centrés (points bleus) sur la ligne centrale de la forme. (c) Les points en entrée (rouge) sont extraits à la surface d’un anévrisme bilobé. Le BM initial est formé d’un seul blob (en blanc). (d) Résultat après 100 subdivisions. La surface implicite (en blanc) modélise le nuage de points (rouge) malgré sa forme complexe.

niveaux entre 0.025 et 0.25 mm d’écart type, chaque forme ayant un rayon typique de 1 mm. Nous avons fixé t_g à trois fois l’écart-type et N_s a pris sa valeur maximale (100) afin d’évaluer la capacité naturelle de LIM à produire des modèles compacts (i.e. avec peu de blobs). Lors de ces expériences, nous avons remarqué que la distance moyenne à la surface de référence n’était pas assez discriminante et nous avons donc opté pour une mesure plus exigeante et robuste, à savoir le 90^e percentile de la distance (seuil de distance sous lequel 90% des 1000 points échantillonnés sur la surface du BM se retrouvent). Nous avons utilisé à la fois la distance euclidienne et son approximation d_T proposée par Taubin [Tau91] (voir Équation 4.16) afin de valider la pertinence de cette dernière. Le score donnant le 90^e percentile de la distance euclidienne est noté d_0 et celui basé sur son approximation d_T est noté d_a .

Ces premières expériences ont permis de confirmer le bon comportement de LIM face au bruit sur les données en entrée, avec des scores d_0 médians de 0.1 mm ou moins, et sa capacité à produire des modèles compacts (nombre médian de blobs entre 9 et 15). La Figure 4.13 donne des exemples de résultat en 2D et en 3D. Le premier met en évidence la capacité de LIM à produire des blobs centrés sur la ligne centrale de la forme, bien qu’aucune contrainte ne l’y incite a priori. Sans être une preuve, c’est un bon indicateur de la qualité modélisatrice de LIM. Le résultat 3D illustre la capacité de LIM de s’ajuster à des formes complexes, telles que celle d’un anévrisme bilobé.

Enfin, l’approximation sur la distance a été validée, les scores d_0 et d_a étant similaires, même si d_a a tendance à sous-estimer la distance pour les points les plus éloignés. Ce dernier point nous a permis de définir un objectif évaluable pour les données patients.

Données patients

Nous avons utilisé les mêmes 10 patients que pour l’évaluation de RBT (voir la Section 4.2.5). En l’absence de méthode de référence, notre objectif a été redéfini comme l’évaluation de la capacité de LIM à ajuster une surface implicite à un nuage de points, sachant que ceux-ci sont bruités (amplitude de l’ordre de la taille voxel), et que l’échantillonnage qu’ils offrent de la surface vasculaire peut présenter des trous. La mesure d_a ayant été validée sur les exemples synthétiques, nous l’avons utilisée comme critère. Notre étude a porté sur les deux paramètres t_g , pris dans $\{0.2, 0.3, 0.5\}$ (en mm), la taille voxel étant de 0.2 mm, et N_s , pris dans $\{30, 50, 100\}$. 9 configurations ont ainsi été comparées tant du point de vue de la compacité des BM finaux, que de la précision. Modéliser l’arbre vasculaire des 10 patients a nécessité 87654 BM. Nos expériences nous ont conduits à clairement préconiser des valeurs de $t_g = 0.3$ mm et $N_s = 100$ pour la 3DRA.

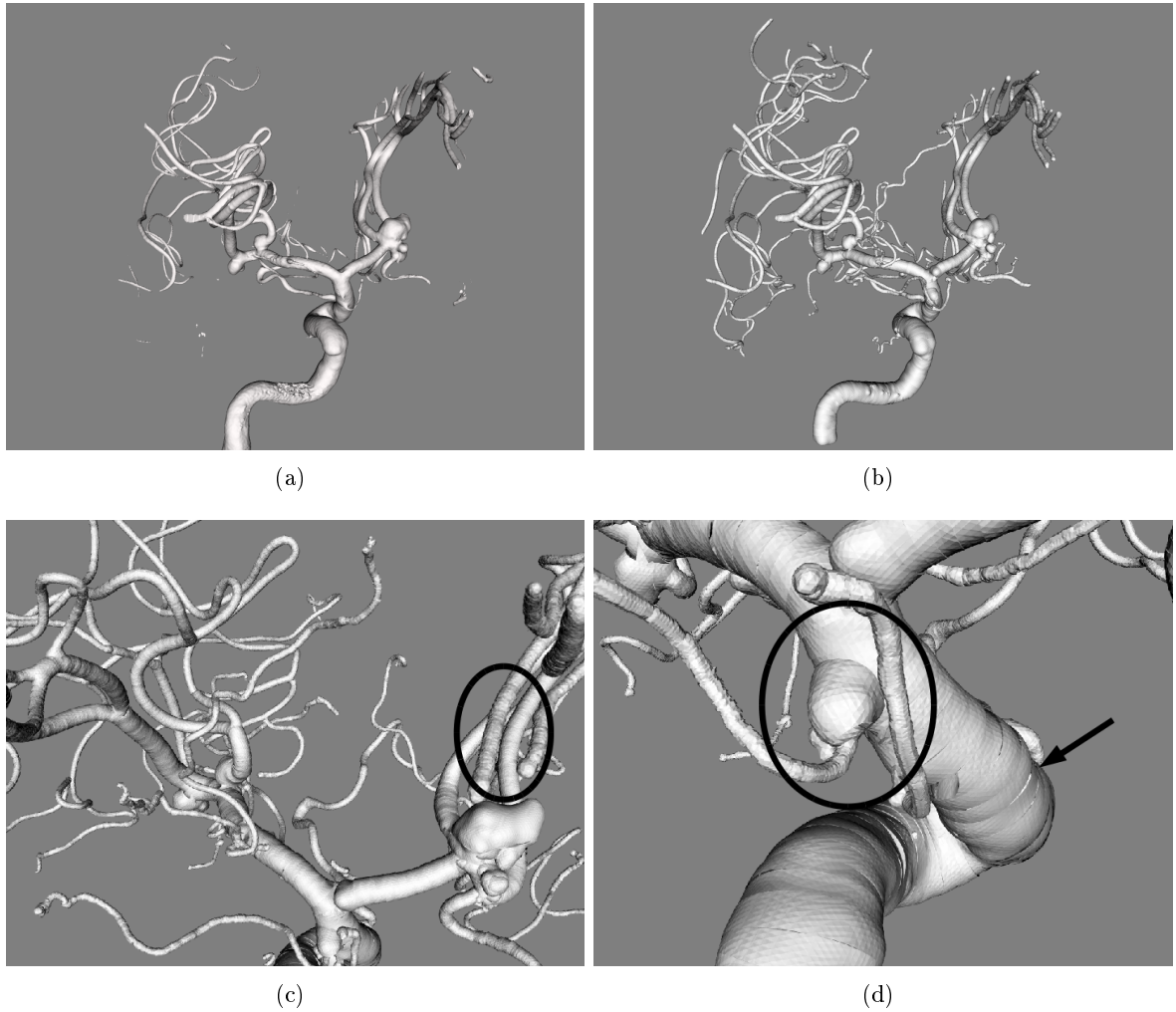


FIGURE 4.14 – Exemple de résultat pour RBT associé à LIM sur données patients. (a) Isosurface obtenue par *Marching Cubes* pour référence. (b) Résultat obtenu par LIM, montrant une bien plus grande richesse et continuité vasculaires. (c) En combinant RBT et LIM le problème du *Kissing Vessel* est correctement géré (encerclé en noir). (d) Exemples de difficultés parfois rencontrées : (encerclé en noir) pour modéliser l'embranchement d'une petite artère (artère ophtalmique) sur une grosse artère (carotide interne) ; (flèche noire) de plus, de petites discontinuités sont encore visibles entre BM adjacents.

Sous ces conditions, 99% des BM s'ajustent à leurs points cibles avec une précision d_a inférieure à la taille voxel. Les BM comportent en moyenne 6.4 blobs. Cela permet une description des surfaces locales avec en moyenne 26 paramètres, là où de nombreux triangles sont nécessaires pour une représentation sous forme de maillage. Selon un calcul médian, il faut 104 fois plus de triangles que de blobs pour représenter une surface vasculaire. La Figure 4.14 fournit un exemple illustratif de nos résultats en associant RBT et LIM, ainsi que quelques difficultés résiduelles.

4.4 Application à la simulation : implantation et validation

L'association de RBT avec LIM fournit des modèles de réseaux vasculaires qui sont en théorie bien adaptés aux calculs réalisés de manière intensive lors d'une simulation. Mais seule une mise en pratique permet de s'assurer de son adéquation et de son impact sur les performances de la simulation.

Nous commençons par rappeler et préciser le contexte clinique et technique dans lequel ces études ont été menées.

Nous reprenons ensuite les détails de l’implantation permettant d’utiliser nos modèles implicites lors d’une simulation, et la comparaison qui en a été faite avec l’usage d’un modèle de surface triangulée. Cette étude a été publiée dans [5], mais de nombreux détails supplémentaires sont rapportés dans [Yur14].

Cependant, la comparaison du même pipeline de simulation avec un autre modèle de surface ne renseigne en rien sur la validité de la simulation elle-même. Les performances actuelles, à la fois en temps de calcul, et au niveau de la complexité des équations physiques simulées, permet aujourd’hui de se poser la question de l’évaluation de la précision de la simulation, en termes quantitatifs. Notre approche de ce point de vue présente deux caractéristiques. D’une part, notre idée consiste à comparer les images simulées avec des images réelles. D’autre part, nous n’avons jamais proprement formalisé cette question. Le texte que nous présentons en fin de section est donc une première tentative de regroupement des différents éléments que nous avons proposés à travers plusieurs articles [17, 24, 26, 52], en vue d’élaborer une réponse plus structurée.

4.4.1 Contexte de la simulation

Notre contexte clinique en général est la neuroradiologie interventionnelle, et en particulier pour cette étude, l’embolisation des anévrysmes par déploiement de coils. Tous les micro-outils partagent donc la forme d’un fil flexible que le neuroradiologue interventionnel pousse, tire et tourne afin de naviguer dans l’arbre vasculaire du patient.

La forme d’un tel fil émane de l’équilibre de forces internes et externes. Les forces internes sont encapsulées dans un modèle à éléments finis représentant le fil comme une séquence de poutres flexibles connectées par leurs extrémités [CDL⁺05]. Les forces externes sont issues du geste du praticien ainsi que de contraintes visant à résoudre la collision avec la surface vasculaire via les lois de contact de Signorini, et de friction de Coulomb [DAK04]. Ce modèle mécanique est implanté dans la plateforme ouverte de simulation interactive Sofa [FDD⁺12]. La gestion des surfaces triangulées y étant déjà présente, nous y avons ajouté celle de nos modèles implicites. Si la valeur des forces externes est différente, puisque liée au modèle de surface, la résolution du système mécanique est la même dans les deux cas.

4.4.2 Implantation de notre modèle dans un contexte simulation

Le modèle mécanique, formé de la succession des poutres, est projeté sur un modèle géométrique polygonal obtenu par échantillonnage de la courbe continue sous-jacente. Lors d’une boucle de simulation, le modèle mécanique commence par évoluer librement, sans prendre en compte la contrainte des vaisseaux sanguins. Puis, chacun des points échantillonnés est passé en revue afin de détecter une possible collision et le cas échéant la gérer en prenant en compte les forces externes ainsi définies pour terminer le pas de simulation [DLC07].

Détection de la collision

La collision d’un point avec une surface implicite se détecte en regardant le signe de la valeur de la fonction implicite en ce point. Dans notre cas, un point a traversé la surface si la fonction du BM local (donnée par f dans l’Équation 4.9) est négative. Nous avons vu dans la section précédente que le calcul de cette fonction est direct et rapide, au vu de la compacité des BM obtenus. Le problème à résoudre est donc celui de la sélection du BM local adapté à la gestion de la collision pour le point. Chaque BM local est associé à un nœud de l’arbre RBT. L’approche naïve consisterait à choisir le nœud dont le centre est le plus proche du point, mais elle n’est pas pertinente.

Tout d’abord, choisir le centre globalement le plus proche pourrait empêcher de détecter la collision. Ce cas est illustré sur la Figure 4.15 où l’extrémité du micro-outil change de vaisseau entre deux temps simulés successifs, sans qu’aucune collision ne puisse être détectée. Par conséquent seuls les nœuds topologiquement voisins du précédent sont considérés pour l’association.

En second lieu, choisir le nœud le plus proche, peut échouer si les rayons associés à ces nœuds sont trop différents. C’est le cas lorsqu’une petite artère se branche sur un gros vaisseau. La

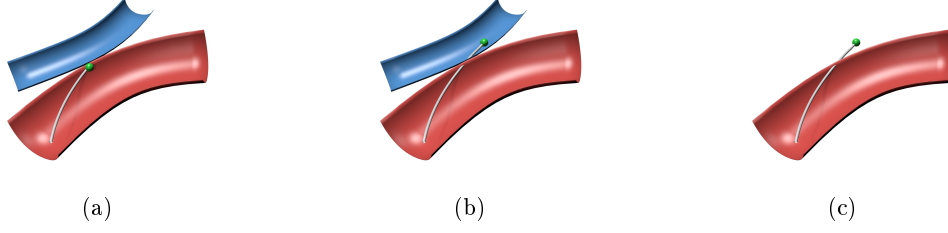


FIGURE 4.15 – Le choix du nœud le plus proche peut causer des artefacts indésirables pendant la simulation, par exemple ici dans un cas de *Kissing Vessel* où deux vaisseaux sanguins (en rouge et en bleu) viennent se frôler. (a) et (b) État de la simulation à deux temps simulés successifs. L’extrémité du cathéter (point vert) passe d’un vaisseau à l’autre. Si l’on associe chaque point à son BM le plus proche, cette transgression passe inaperçue (la fonction implicite associée renvoie à chaque fois une valeur positive). (c) En revanche, ne prendre en compte que les BM topologiquement voisins permet de s’affranchir du vaisseau bleu et détecter correctement la collision.

Figure 4.16 schématise ce cas. Choisir le nœud le plus proche revient à déterminer dans quelle cellule se trouve le point, dans le diagramme de Voronoï des centres des nœuds. Dans le cas évoqué ici, la frontière entre petite artère et gros vaisseau n’est pas bien localisée et des erreurs de détection de collision peuvent se produire. Le diagramme de puissance, permet de remédier à cela en adaptant la taille des cellules à leur rayon estimé par RBT. Ce diagramme est généré par la pseudo-distance :

$$d^*(P, C) = |P - C|^2 - r^2 \quad (4.19)$$

où P est le point considéré sur le micro-outil, C est le centre du nœud et r son rayon.

À chaque pas de simulation, chaque point de l’outil est traité indépendamment. On lui associe le nœud le plus proche au sens de d^* , parmi les voisins topologiques du nœud associé au temps précédent. Notre approche est illustrée sur la Figure 4.17 (reprenant la Figure 4.1) qui montre le BM sélectionné pour le point extrémité du cathéter à différents moments d’une navigation simulée dans un arbre vasculaire de patient.

Résolution des collisions

La gestion d’une collision pour un point P implique la gestion du contact et celle de la friction. Concernant le contact, il se résout en projetant P sur la surface vasculaire en P' . D’après la loi de Signorini, la force de réponse au contact est alors le vecteur d’interpénétration $P' - P$. La projection se fait par quelques itérations de Newton-Raphson. Si f est la fonction implicite, il est à noter qu’une telle itération déplace le point de $d_T(P) = \frac{|f(P)|}{|\nabla f(P)|}$ le long du vecteur gradient normalisé $\frac{\nabla f(P)}{|\nabla f(P)|}$. Son calcul est donc rapide puisque nous disposons de formules directes et compactes pour f et son gradient en tout point P .

La même distance peut être exploitée afin de gérer la force de friction. Celle-ci demande de définir le premier point de contact Q_0 , et intègre les déplacements successifs de P' tant que le point P reste en contact. Lors de la première détection d’une collision, Q_0 est le point où la surface a été traversée. Supposons qu’au temps t , le point P_t n’est pas en collision, mais qu’il s’y retrouve au temps suivant $t + 1$, en P_{t+1} . Le point Q_0 est le point du segment $[P_t P_{t+1}]$ où la fonction implicite s’annule. Nous trouvons ce point en déterminant le zéro de la fonction $g(\lambda) = d_T((1 - \lambda)P_t + \lambda P_{t+1})$ par la méthode de la fausse position. La Figure 4.18 illustre cette technique.

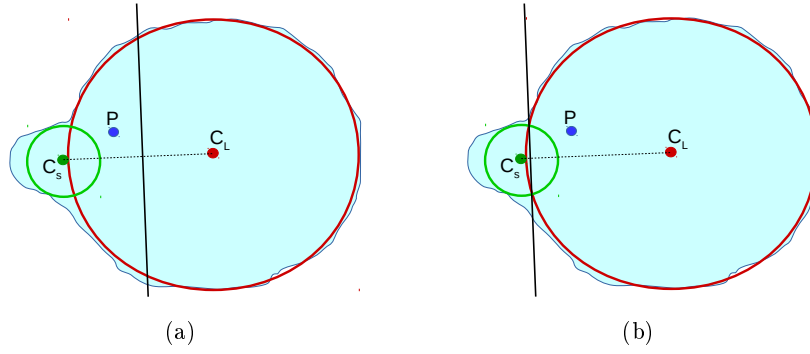


FIGURE 4.16 – Cas de l’embranchement d’une petite artère sur un gros vaisseau. Le nœud modélisant la grosse artère est en rouge, de centre C_L et de grand rayon. Il a pour voisin topologique le nœud vert, de centre C_s et de petit rayon, qui modélise la partie proximale de la petite artère. Dans le cas (a), le point P est associé au nœud le plus proche, de même centre que la cellule de Voronoï dans laquelle il se trouve : se trouvant du même côté que C_s par rapport à la médiatrice (ligne noire continue) du segment reliant les deux centres (ligne noire en pointillés), c’est le nœud vert qui est choisi, et le point P est détecté en collision. Dans le cas (b), les cellules sont celles du diagramme de puissance : la ligne de séparation est déplacée vers C_s et P est correctement associé au nœud rouge.

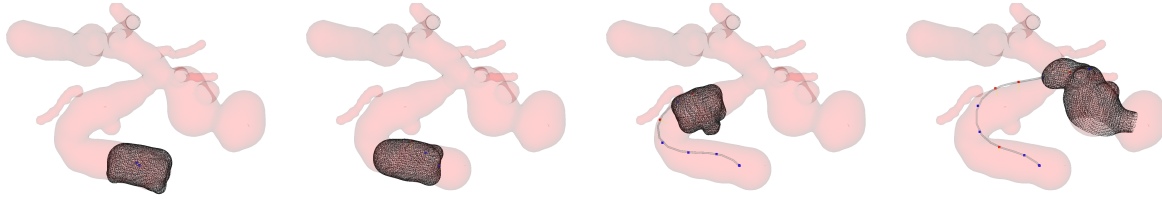


FIGURE 4.17 – Illustration de notre procédure de sélection du BM local pour le point extrémité du cathéter. De gauche à droite, différents temps de la simulation sont montrés chronologiquement lors de la progression du cathéter dans l’arbre artériel issu de données patients. La surface du BM sélectionné est affichée sous la forme d’une surface triangulée, uniquement pour les besoins de la visualisation.

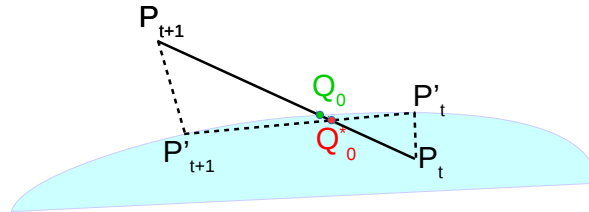


FIGURE 4.18 – Méthode de la fausse position pour approximer le point de contact Q_0 par le point Q_0^* . L’intérieur de la surface implicite est en bleu ciel. Au temps t , le point P_t n’est pas en collision. Au temps suivant $t+1$, il l’est en sa position P_{t+1} . P'_t et P'_{t+1} sont leurs projections sur la surface, approchées par une itération de Newton-Raphson. La méthode de la fausse position suppose que les droites $(P_t P'_t)$ et $(P_{t+1} P'_{t+1})$ sont parallèles. Le théorème de Thalès permet donc de déterminer $Q_0^* = (1 - \lambda_0)P_t + \lambda_0 P_{t+1}$ avec $\lambda_0 = \frac{|P'_t - P_t|}{|P'_t - P_t| + |P'_{t+1} - P_{t+1}|}$.

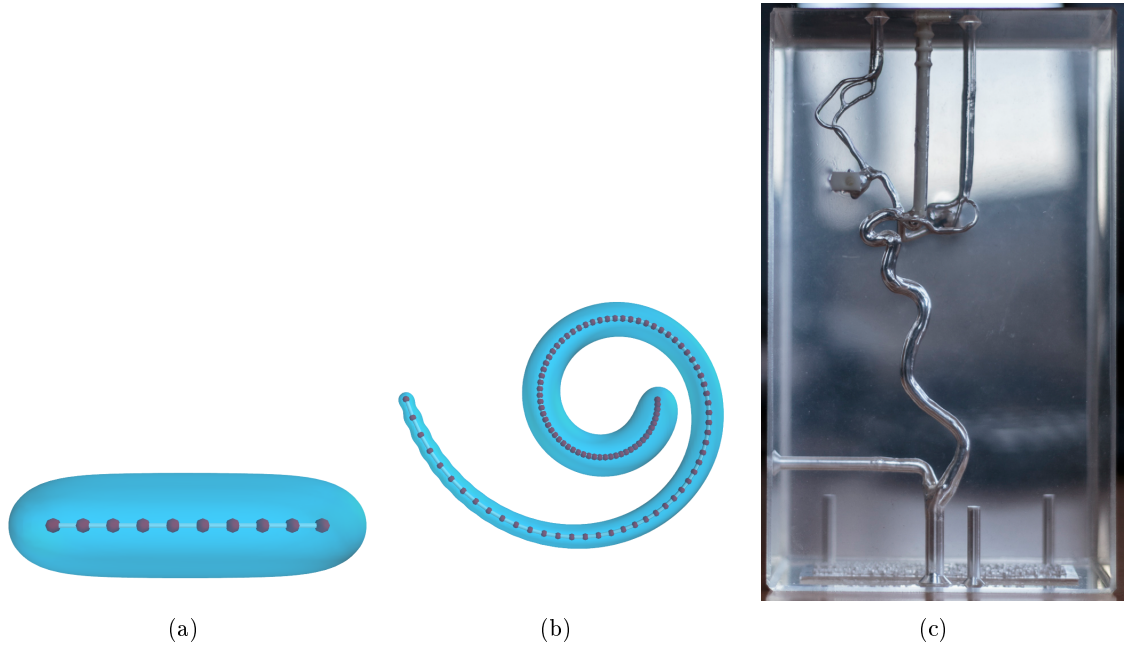


FIGURE 4.19 – Données utilisées pour valider notre modèle implicite lors d’une simulation. (a) capsule, forme la plus simple de vaisseau ; (b) spirale, exigeante pour sa gestion de contacts multiples ; et (c) fantôme vasculaire en silicone, pour un cas plus réaliste.

Évaluation

L’impact de notre modèle implicite sur la simulation a été évalué en regard de trois bénéfices attendus : temps réduit pour gérer les collisions, meilleur réalisme, and meilleure stabilité numérique. Nous avons utilisé deux formes synthétiques (une capsule, et une spirale, cette dernière à trois niveaux de résolution différents pour le maillage) et les données obtenues sur un fantôme vasculaire en silicone (Elastrat, Genève). Une illustration en est donnée sur la Figure 4.19 et les détails de leur génération et du paramétrage de la simulation sont disponibles dans [5].

Le temps moyen mesuré pour gérer les collisions est plus petit par deux ordres de grandeur avec notre modèle qu’avec une surface triangulée, pour un nombre comparable de points en contact. Un autre avantage d’employer des modèles locaux est que ce temps ne dépend pas de la complexité globale de la surface. Enfin, la sélection du BM local se fait en temps constant puisqu’elle implique seulement les BM dans un voisinage topologique du précédent.

La Figure 4.20 illustre parfaitement la différence de réalisme visuel entre une surface triangulée et notre modèle implicite. Dans le cas triangulé, même avec un coefficient de friction mis à zéro, l’extrémité du cathéter reste bloquée dans la zone de forte courbure, simulant une friction importante qui n’a pas lieu d’être. Le phénomène, dû à la discrétisation de la surface, s’amointrit en augmentant la résolution, mais le temps de résolution devient prohibitif. Dans le cas implicite, le comportement du cathéter est réaliste, et l’effet de friction recherché à travers une petite valeur du coefficient de friction est bien reproduit.

Enfin, l’amélioration de la stabilité numérique apportée par LIM apparaît avec la spirale. Comme l’illustre la figure 4.21, le cathéter, qui reste à l’intérieur de la forme avec LIM, la traverse systématiquement en utilisant une surface triangulée dès lors que le mouvement appliqué à l’outil est un peu rapide. Une première raison de ce phénomène est liée au fait que la simulation, pour rester interactive, n’atteint pas la position statique à la fin de chaque pas de temps, mais doit compter avec des imprécisions dues au régime dynamique. Ces imprécisions sont pires dans le cas triangulé car le temps de résolution des contraintes est plus long. Une deuxième raison est que la définition continue et lisse du gradient à la surface dans le cas implicite conduit à des valeurs des forces de contact plus cohérentes le long de la surface, et donc une meilleure stabilité numérique.

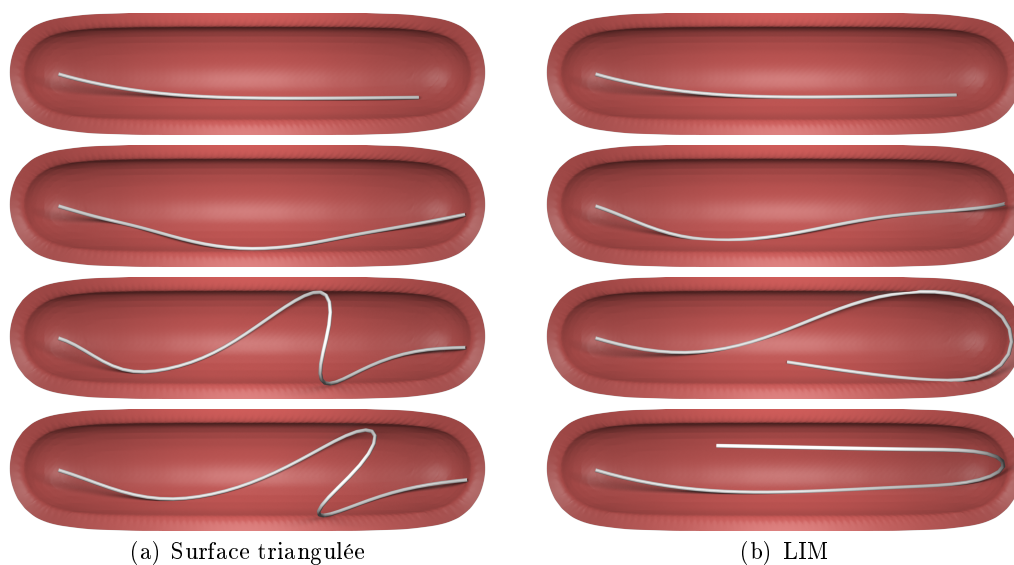


FIGURE 4.20 – Simulation de navigation de cathéter dans la forme capsule, comme surface triangulée (à gauche), et comme surface implicite LIM (droite). Même avec une résolution très fine (12k triangles), un phénomène de friction apparaît avec la surface triangulée. Cette friction est artificielle, et non réaliste, car le coefficient de friction utilisé est nul. LIM (12k blobs) reproduit en revanche un cathéter qui glisse de manière réaliste dans la forme (avec un petit coefficient de friction).

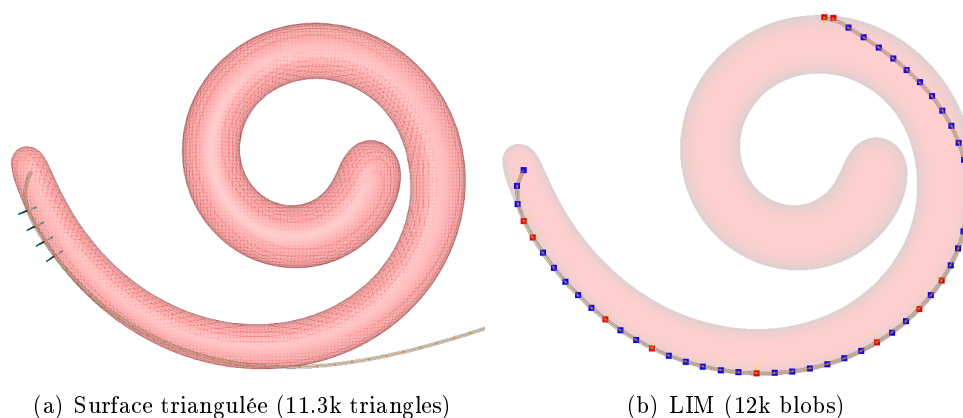


FIGURE 4.21 – Illustration de la meilleure stabilité numérique avec LIM. (a) Le cathéter traverse systématiquement la surface triangulée dès que les forces de contact deviennent trop importantes. (b) Le cathéter suit un chemin réaliste, sans provoquer d'échec dans la simulation. À noter que le temps moyen pour gérer les collisions est de 1.84 ms dans le cas triangulé, contre 0.0219 ms avec LIM.

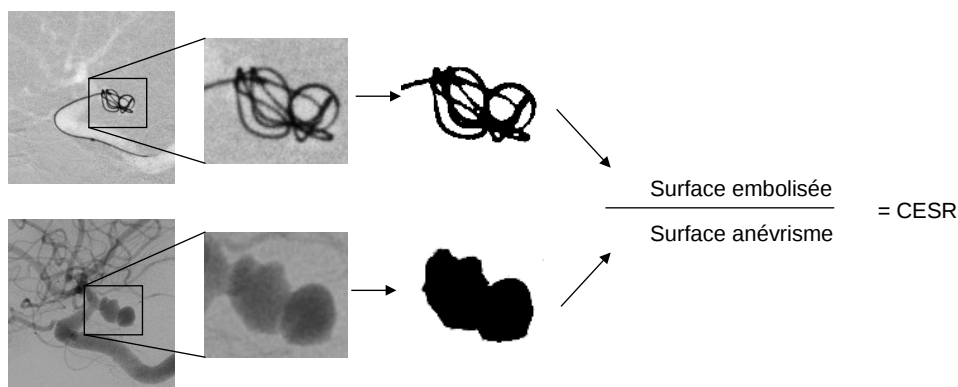


FIGURE 4.22 – Exemple de calcul du CESR (*Coil Embolization Surface Ratio*) pour le cas réel : un masque binaire est calculé pour l'anévrisme dans une image d'angiographie soustraite (bas) et pour le coil pendant son déploiement sous fluoroscopie (haut). Une surface est ainsi calculée dans chaque cas. Le rapport des deux surfaces fournit le CESR. Un calcul semblable est fait pour le cas simulé, dans la même incidence de travail.

4.4.3 Quelques éléments pour la validation d'une simulation interactive

Les résultats de ce travail permettent d'envisager que l'usage du simulateur sera bénéfique pendant la formation initiale et la formation continue des neuroradiologues interventionnels. Mais ils nous incitent aussi à rechercher des performances encore plus grandes afin de pouvoir bénéficier d'une simulation fiable pour l'étape suivante, à savoir assister le médecin pour son planning pré-opératoire.

Cette question nous force à aborder l'évaluation d'une simulation d'un point de vue quantitatif. Nous savons d'avance que toute simulation s'écartera de la situation réelle, ne serait-ce qu'à cause des approximations numériques (linéarisation des équations) ou de modélisation (pas de plasticité des micro-coils, pas de simulation du flux sanguin, ...). Cette différence est acceptable pour les médecins puisque dans la réalité, la répétabilité d'un geste n'est pas absolue. L'objectif du planning pré-opératoire est cependant de comparer différentes stratégies opératoires, soit pour choisir entre différents outils envisageables, soit pour affiner progressivement un geste. Dans les deux cas, le planning est validé s'il conduit, éventuellement au sens statistique, à un succès opératoire, c'est-à-dire à des faibles taux de mortalité et de morbidité qui ne peuvent se mesurer qu'à moyen voire long terme. Les médecins ont bien évidemment développé des critères prédictifs de ce succès, évaluables pendant l'opération. À côté des critères empiriques, relevant souvent du savoir-faire, les sociétés savantes de médecine préconisent aujourd'hui des critères quantitatifs que des études ont pu corrélérer au succès opératoire. Dans le cas de l'embolisation des anévrismes par coils, le CER (*Coil Embolization Ratio*) [SSMS98, TIA⁺02] est par exemple un des facteurs prédictifs d'une bonne stabilité de l'embolisation : une valeur de CER au-dessus de 20% (le CER maximum est 33%) est un bon indicateur d'une absence de recanalisation à un an.

Simuler pour choisir le premier coil

Une étape cruciale dans l'embolisation des anévrismes est le choix et la pose du premier coil : celui-ci doit permet de former, dans le volume de la poche anévrismale, un grillage tridimensionnel complexe, dont les spires s'appuient régulièrement sur la surface de l'anévrisme, certaines traversant le collet, sans toutefois pénétrer le lumen artériel [CJT⁺00]. Notre objectif était d'évaluer la capacité de la simulation à prédire un placement correct du premier coil. Nous avons pour cela proposé dans [24] un critère mesurant le pourcentage de surface anévrismale recouverte par le coil. Nous inspirant du CER, nous avons nommé ce critère CESR (*Coil Embolization Surface Ratio*). Comme illustré dans la Figure 4.22, sa mesure se fait dans la vue de travail : d'une part car nous n'avons pas de reconstruction 3D de ce premier coil pour servir de vérité terrain,

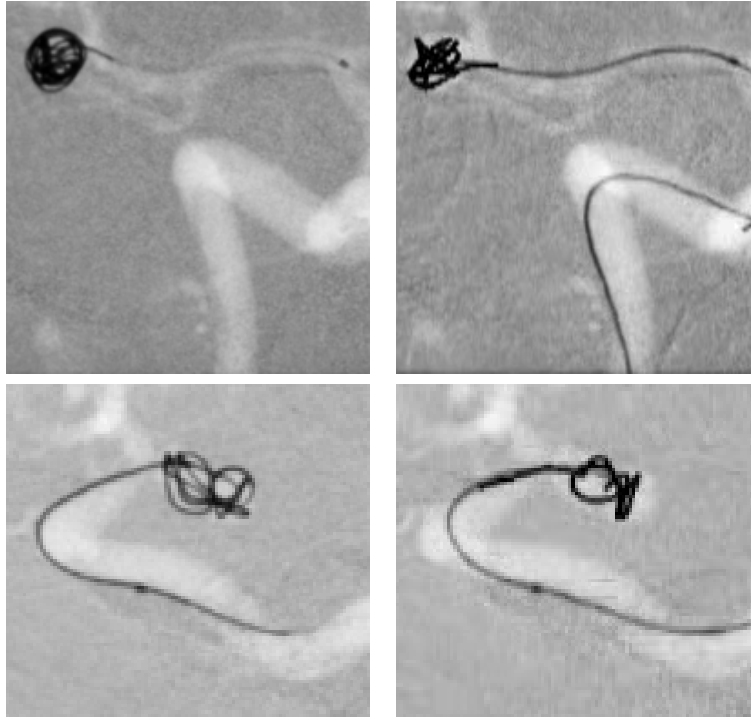


FIGURE 4.23 – Comparaison de déploiements de coils 3D sur des cas réels (colonne de gauche) et simulés (colonne de droite). Un patient différent par ligne. (en haut) CESR réel = 81% vs simulé = 68%. (en bas) CESR réel=24% vs simulé = 17%.

mais également car ce calcul se rapproche de la pratique du neuroradiologue interventionnel qui procède à ce genre d'évaluation en continu pendant le déploiement du coil sous fluoroscopie.

Des exemples de comparaison entre des cas réels et simulés sont présentés sur la Figure 4.23. Ces cas ont été choisis car un coil de forme hélicoïdale classique a été jugé inadapté par les neuroradiologues interventionnels qui ont embolisé l'anévrisme dans chaque cas avec un coil 3D. La mesure du CESR sur les cas simulés a clairement confirmé cette inadaptation du coil hélicoïdal et favorisé le coil 3D, même si le CESR reste sous-estimé par la simulation [24, 52].

Vers une validation de la forme simulée

Ces premières investigations nous ont convaincus de l'importance de mettre au point un critère permettant de valider la forme simulée. Cette conviction a été renforcée par les expérimentations que nous avons menées pour estimer les paramètres mécaniques d'un coil [26], et qui ont montré qu'une fois correctement paramétré, l'algorithme de simulation avait la capacité de reproduire la forme du coil en 3D, aux limitations de l'étude près, celles-ci étant que cette forme était simplement paramétrée (coil hélicoïdal), que le déploiement se faisait sans contact, et que l'étude portait sur le comportement statique du coil.

En collaboration avec l'équipe DEFROST d'Inria Lille, nous avons donc réalisé un premier prototype de banc de test pour la simulation des micro-outils en cours de navigation dans un environnement avec contacts. L'un des enjeux était de pouvoir capturer en 3D des changements de forme potentiellement très rapides auxquels le micro-cathéter est soumis, et que les médecins décrivent comme des «mouvements de fouet». Ce phénomène de «coller-glisser» (*stick and slip transition*) est dû aux frottements et se manifeste quand une zone de contact devient bloquante. Sous l'effet de la force de poussée du manipulateur, l'énergie s'y accumule jusqu'à un seuil où le blocage se rompt, provoquant un changement rapide de forme pour le micro-cathéter. Ce système comporte deux composants. D'un côté, un système de capture du mouvement du cathéter, conçu par DEFROST, permet d'enregistrer avec précision le geste du manipulateur. De l'autre, nous avons développé un système d'acquisition stéréovision rapide, associé à un algorithme de recons-

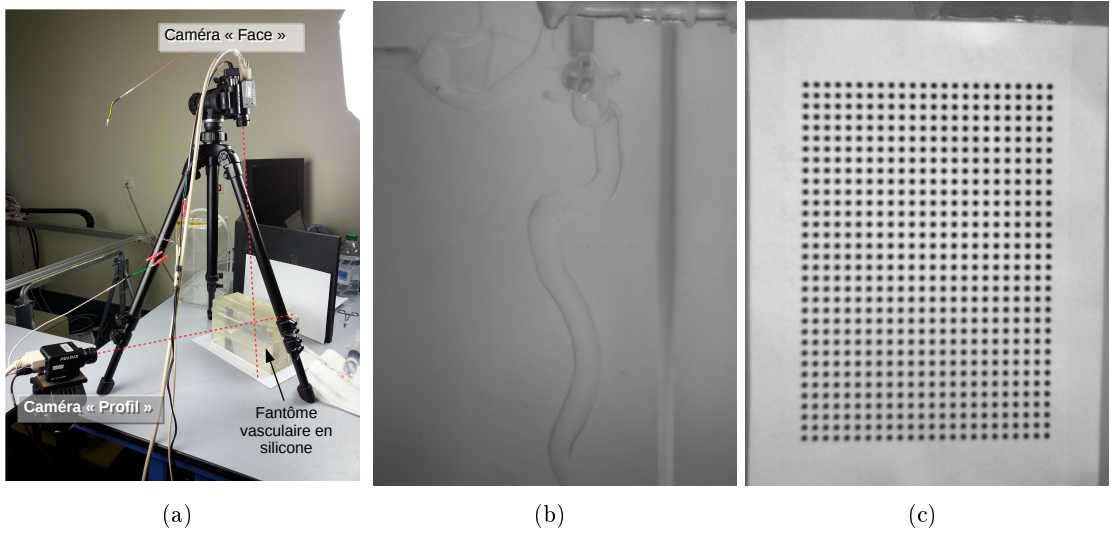


FIGURE 4.24 – Éléments du dispositif expérimental de validation de la simulation. (a) Système de caméras en stéréovision. Les axes optiques (en pointillés rouge) sont quasiment orthogonaux. (b) Exemple d’image du fantôme rempli de liquide de viscosité semblable à celle de sang. (c) Mire de calibration employée pour estimer l’équation du plan de réfraction.

truction 3D de la forme du micro-outil. Un aperçu de leurs caractéristiques, de leur validation et des premiers résultats sont donnés dans [17]. Nous nous focalisons ici sur le système stéréovision.

Les transitions «coller-glisser» sont impossibles à capturer par rayons X, d’une part à cause de la fréquence trop faible (30 images par seconde au mieux) et d’autre part à cause du filtrage temporel appliqué aux images de fluoroscopie afin de les rendre lisibles, mais qui génère du flou cinétique. Nous avons donc exploité le système d’acquisition stéréovision développé pour notre projet d’analyse de la parole qui permet d’acquérir des images de taille 640×480 en niveaux de gris, synchronisées par trigger à 198 images par seconde, soit un intervalle de 5.05 millièmes de seconde entre deux images. (voir Chapitre 2 et particulièrement la Section 2.2.4). La zone d’intérêt étant petite, et l’objet étant filaire, nous avons pu les positionner de telle manière que les axes optiques soient orthogonaux, ainsi que l’illustre la Figure 4.24(a). Cette paire de caméras stéréo a été calibrée grâce à la méthode par damier implantée dans OpenCV¹². Par ailleurs, afin de pouvoir étudier les contacts, nous avons tiré profit du fantôme vasculaire déjà utilisé pour valider LIM (voir Figure 4.19(c)). Ce fantôme est transparent, ce qui permet une bonne visualisation du micro-cathéter en cours de navigation. Le silicone est cependant un matériau réfractant, dont nous connaissons l’indice de réfraction $n_s = 1.43$ (donnée constructeur). Il se trouve que cet indice est quasiment le même que celui du produit visqueux injecté dans le fantôme pour simuler le sang durant la navigation. Nous pouvions donc considérer que le fantôme était un bloc homogène de silicone (voir Figure 4.24(b)). Enfin, les caméras étant en positions orthogonales, nous avons placé le fantôme de telle manière que tous les rayons reçus par une caméra donnée étaient réfractés par un seul et même plan. L’équation de chaque plan est calibrée à partir d’une mire planaire (Figure 4.24(c)).

Qu’il soit à l’air libre ou dans le fantôme, la segmentation du cathéter dans les images peut se faire par simple seuillage associé à un léger filtrage gaussien de lissage (voir des exemples sur la Figure 4.25). À l’air libre, la reconstruction se fait simplement par triangulation, en faisant l’hypothèse que l’extrémité distale du cathéter peut être appariée dans les deux vues : cette extrémité est reconstruite, puis on procède par propagation de la reconstruction le long de la forme. Un filtrage gaussien 3D permet de produire une forme régularisée. Le milieu réfractant du fantôme rend la triangulation plus délicate. Sans perte de généralité, considérons un pixel dans l’image de face. La caméra de face étant calibrée, il est aisé de déterminer l’équation du rayon

12. http://docs.opencv.org/3.2.0/dc/dbb/tutorial_py_calibration.html

vecteur associé à ce pixel. De plus, l'équation du plan de réfraction étant connue, il est tout aussi aisé de déterminer le point d'intersection entre ce rayon vecteur et le plan de réfraction. L'application de la loi de Snells-Descartes conduit alors à l'équation du rayon réfracté dans le silicone. La difficulté de la triangulation provient du fait que ce rayon, segment droit, se projette sur la vue de profil, réfracté selon une quartique [CS09]. Nous discrétisons cette quartique pour trouver son intersection avec la forme segmentée dans l'image de profil et procéder ensuite par triangulation des points ainsi mis en correspondance. Une étude préliminaire a permis d'estimer l'écart-type de cette reconstruction en regard du bruit dans les images (statistiques sur 1000 images statiques du cathéter) à $[0.0075, 0.0175, 0.0457]$ mm selon chaque axe X,Y,Z (les deux derniers étant approximativement alignés avec les axes optiques des caméras).

Un exemple de reconstruction lors d'un mouvement brusque de «coller-glisser» est présenté sur la Figure 4.25. L'extrémité du cathéter se déplace de 10.1 mm en l'espace de 35.35 millièmes de secondes (8 images), soit une vitesse moyenne de 285 mm/s. Sur cette séquence, les vitesses instantanées, mesurées entre deux images consécutives, donne le profil indiqué sur la Figure 4.25. Une vitesse maximale de presque 500 mm/s a ainsi pu être mesurée.

4.5 Conclusion et perspectives

Ce chapitre reprend nos travaux en modélisation vasculaire. Notre point de vue est assez original par le contexte applicatif de simulation, et en particulier de la gestion de collisions.

Nous avons proposé un nouveau modèle formé d'un arbre de fonctions implicites. La topologie de l'arbre est extraite grâce à un algorithme de suivi des vaisseaux, nommé RBT, dont nous avons prouvé l'excellente performance sur des arbres vasculaires intra-cérébraux. La particularité de cet algorithme est aussi de fournir des ensembles de points locaux, qui échantillonnent de manière robuste la surface vasculaire dans le voisinage de chaque point de la ligne centrale des vaisseaux sanguins.

Notre algorithme LIM est inspiré de travaux assez anciens [Mur91, TBC95, BTG95]. Il minimise une fonctionnelle d'énergie adaptée aux BM et à des ensembles de points pouvant comporter des trous, suivant une heuristique dont l'efficacité a été prouvée sur des cas synthétiques et cliniques. Le résultat est une amélioration significative de la performance, en termes de précision et d'automatisation, qui nous a permis de modéliser des arbres vasculaires entiers avec très peu d'intervention manuelle.

Nos travaux actuels visent à améliorer cette modélisation suivant deux axes. Le premier, à relativement court terme, vise à simplifier encore l'arbre produit. Cet arbre contient en effet un grand nombre de points centraux, qui sont très rapprochés, notamment dans les petits vaisseaux : la hauteur d'un cylindre, et donc le pas de suivi, est en effet lié au rayon couramment estimé pour le vaisseau. Les BM ne sont alors pas assez «longs» pour rester valides entre deux pas de temps si le cathéter avance trop vite, ce qui peut générer des fausses collisions. Nous avons appliqué un simple rééchantillonnage des lignes centrales, qui préserve les points de bifurcations. Le nombre de BM passe de quelques milliers à quelques centaines. Les ensembles de points locaux liés à chaque nouveau centre sont créés par concaténation. Cependant, les BM estimés par la suite sont moins compacts, mais surtout, la continuité entre BM adjacents est plus difficile à assurer. Une contrainte explicite doit être mise en œuvre. Nous envisageons d'imposer que les BM adjacents partagent des blobs communs. La conséquence est une perte d'indépendance de l'optimisation de chaque BM, avec un impact sur le temps de modélisation. Dans l'hypothèse d'une minimisation basée sur une descente de gradient, l'impact sur la matrice jacobienne de l'énergie est que celle-ci, de diagonale par blocs devient tridiagonale par blocs, ou plutôt diagonale par blocs recouvrants. Son inversion pourrait donc rester efficace. Le deuxième axe, à plus long terme, cherchera à s'affranchir des points sélectionnés par RBT à la surface locale du vaisseau, et que LIM approxime par une surface implicite. Ceci pour deux raisons. Tout d'abord, un retour à l'image permettrait de prendre en compte des contours de vaisseaux mal échantillonnés par RBT, en particulier au niveau des zones pathologiques qui s'écartent trop de l'hypothèse de forme cylindrique pour le vaisseau. Notre idée est de modifier le terme d'énergie \mathcal{E}_d , pour le remplacer

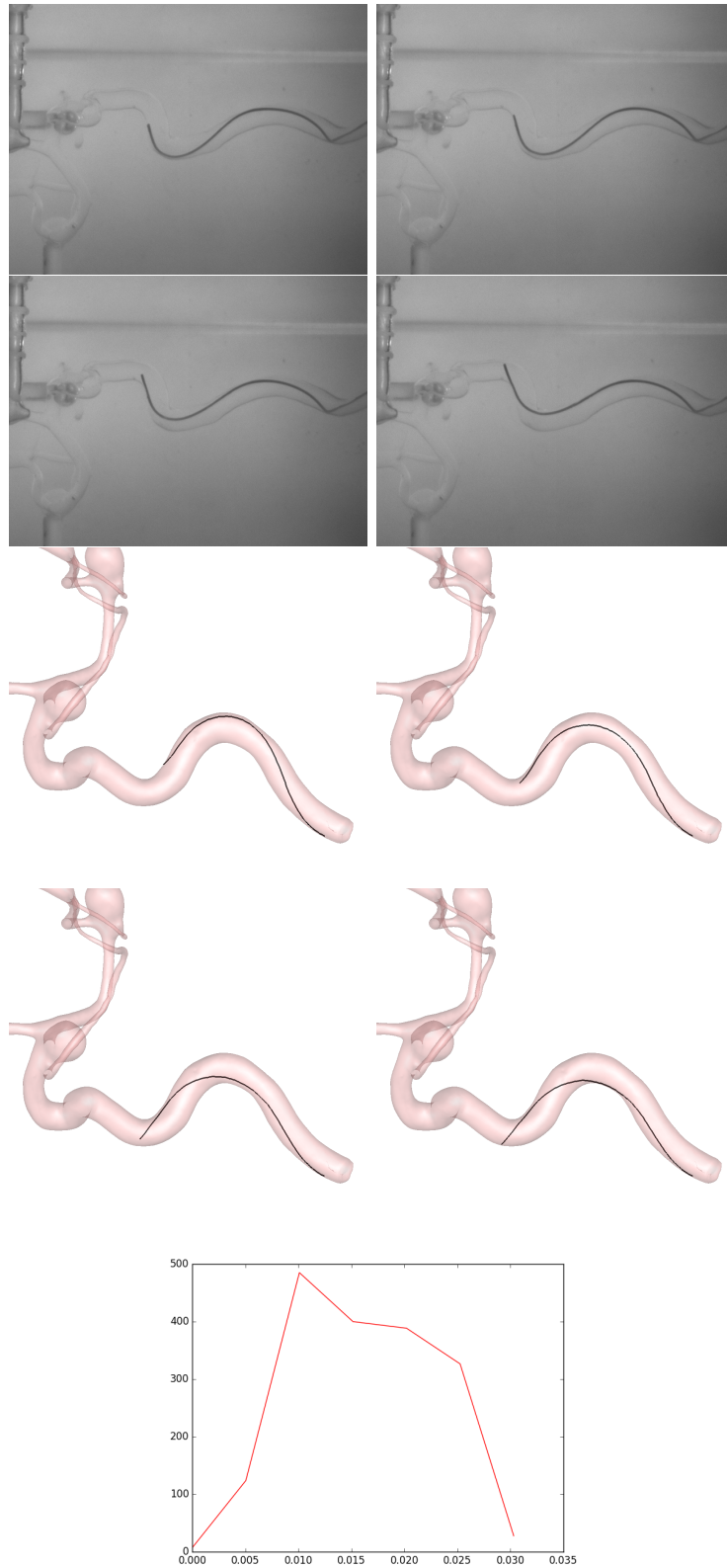


FIGURE 4.25 – Reconstruction lors d’une transition «coller-glisser». En haut, séquence capturée par la caméra de face (une image sur deux est montrée dans la séquence par souci de concision, écart entre deux images=environ 10 millièmes de seconde). En dessous, reconstruction 3D, visualisée selon une incidence oblique. En bas, profil de vitesse instantanée estimée entre deux images successives. Une vitesse maximale de 485.13 mm/s a pu être ainsi mesurée.

par l'intégrale de la norme du gradient image sur la surface implicite courante. Le deuxième point est plus technique et se réfère au terme d'énergie \mathcal{E}_a visant à minimiser la courbure moyenne : l'Équation 4.13 montre que la courbure est calculée sur les points donnés en entrée alors qu'elle devrait être intégrée sur la surface implicite courante. Ces deux axes peuvent donc être abordés par la question du calcul d'une intégrale sur la surface implicite. Une voie à explorer est l'approche variationnelle comme celle proposée par Slabaugh et al. pour des RBF [SDU07]. Une autre voie consiste à procéder par intégration de Monte-Carlo, en échantillonnant la surface implicite, par exemple en utilisant [WH94].

Une deuxième direction de recherche vise à parfaire le banc de test permettant de valider une simulation en regard de données réelles. D'autres équipes de recherche poursuivent un objectif similaire [DNH16], et proposent un système mieux validé et caractérisé, mais qui ne peut pas gérer les situations de collision réalistes. Nous souhaitons donc mieux caractériser expérimentalement notre système et en améliorer la robustesse en exploitant les résultats de recherche en reconstruction biplan de cathéter que nous avons décrits dans le Chapitre 3.

Chapitre 5

Conclusions et perspectives

Sommaire

5.1	Conclusion générale	113
5.2	Vers une simulation guidée par l'image	113
5.2.1	Détecter	114
5.2.2	Prédire	115
5.2.3	Adapter	116

5.1 Conclusion générale

Au travers d'une sélection de mes travaux de recherche, les chapitres qui constituent ce mémoire en dessinent l'objectif générique comme l'élaboration de modèles, géométriques, statistiques, physiques, à partir de données qui sont principalement des images. Ces modèles doivent pouvoir représenter des objets déformables, comme les micro-outils utilisés en neuroradiologie interventionnelle ou des organes, auquel cas il leur faut en plus pouvoir s'adapter aux morphologies des différents patients ou locuteurs. Cet objectif a été illustré dans les deux contextes applicatifs que sont l'analyse articulaire et la neuroradiologie interventionnelle. Ces exemples ont permis de mettre en évidence deux caractéristiques de ma démarche. D'une part, l'approche est pluridisciplinaire, impliquant une collaboration étroite avec des médecins, industriels, ou des spécialistes d'autres domaines (analyse de la parole, modélisation physique). D'autre part, les thèmes abordés couvrent la chaîne d'acquisition, de traitement, et d'analyse des données jusqu'à la construction du modèle. Cette importance de l'application dans le pilotage de mes recherches implique un effort particulier d'assimilation de la problématique posée par le tiers collaborateur et de sa pratique professionnelle, ainsi qu'une exigence constante de validation des solutions développées.

J'ai choisi de livrer en fin de chaque chapitre les extensions me semblant en continuité naturelle avec le travail qui y était présenté, afin que leur appréhension bénéficie de la prégnance du contexte applicatif chez le lecteur. Les pistes évoquées seront explorées à court ou moyen terme. Je discute dans le reste de ce chapitre de conclusion un axe de recherche que je n'ai qu'effleuré pour le moment mais qui me paraît très prometteur et que je souhaite développer à plus long terme.

5.2 Vers une simulation guidée par l'image

Dans l'introduction de ce mémoire, nous énoncions notre but comme étant de «proposer un environnement où le chirurgien ou le radiologue interventionnel pourra à nouveau concevoir, planifier, répéter et réaliser son geste en 3D». Les modèles mécaniques, et plus globalement physiques, nous paraissent très prometteurs pour l'atteindre par leur capacité à prévoir et à expliquer les phénomènes observés. Un exemple illustratif est celui que nous avons donné dans [52].

Il concerne le déploiement d'un coil dans un anévrisme. Le choix du premier coil se fait sur un critère géométrique : le grillage formé en 3D doit couvrir l'ensemble de la poche anévrismale, sans qu'aucune boucle ne dépasse dans la lumière de l'artère parente. Nous avons montré que la simulation physique permettait aussi d'estimer les forces de pression exercées par le coil sur la paroi anévrismale, forces qui pourraient expliquer des ruptures de la poche, ou plus généralement prédire une évolution défavorable liée à une distension de la poche anévrismale et donc une recanalisation, que l'aspect visuel seul ne saurait prévoir.

Cependant, force est de constater aujourd'hui que la simulation physique, bien que de plus en plus performante et interactive, s'éloigne très rapidement de la réalité terrain, par exemple si on la confronte à une image interventionnelle. Les raisons en sont multiples et aucun maillon de la chaîne n'est épargné : les modèles physiques sont par nature des approximations, dont la fidélité dépend de plus de paramètres délicats à fixer ; les modèles géométriques sont construits à partir de données bruitées et à résolution finie ; les schémas numériques sont issus d'un compromis entre précision et temps de calcul, sous contrainte de leur stabilité ; les capteurs utilisés pour fixer la commande ou pour vérifier la validité de la simulation produisent également des données bruitées, voire artéfactées et limitées en résolution spatiale et temporelle. Ces raisons nous font penser aujourd'hui qu'aucune modélisation a priori ne permettra d'atteindre une fidélité de simulation telle qu'elle peut s'affranchir des images interventionnelles.

En revanche, il nous semble tout à fait envisageable de coupler dynamiquement la simulation et les données images, réalisant ainsi une simulation guidée par l'image, de telle manière que le résultat de la simulation est confirmé en permanence par l'image. Les applications sont diverses mais nous en avons identifié deux qui servent de support à notre approche. La première est la reconstruction monoculaire basée modèle. La simulation étant réalisée en 3D, le pari est d'affirmer que le modèle physique est suffisamment complet et contraint pour lever les ambiguïtés des données, notamment dans l'axe de profondeur du champ visuel. La seconde est la réduction de la dose de rayons X par réduction de la fréquence d'acquisition de la fluoroscopie. La qualité de prédiction de la simulation permettrait d'extrapoler les données images afin de conserver une fréquence d'affichage de 15 ou 30 images par seconde, tout en n'acquérant qu'une image toutes les secondes par exemple. C'est l'objectif applicatif de la thèse de Raffaella Trivisonne (débutée en septembre 2015, dirigée par Stéphane Cotin et que je co-encadre). Le contexte plus général est celui des Systèmes Applicatifs Guidés par des Données Dynamiques (DDDAS pour *Dynamic Data Driven Applications Systems*, concept développé par la *National Science Foundation* en 2004 [Dar04]). La boucle de rétro-contrôle envisagée se fonde sur trois tâches : détecter, prédire et adapter.

5.2.1 Détecter

La détection est celle du changement et elle prend deux sens.

En premier lieu, détecter signifie capturer le mouvement, autrement dit mesurer la commande à appliquer au système simulé. Dans le cas de la neuroradiologie interventionnelle, nous avons décrit un dispositif de capture du geste du praticien dans [17] et nous sommes impliqués dans un projet avec le CIMIT (Mark Ottensmeyer, Boston) financé par le département de la défense américain, dont un objectif central est le développement d'un tel dispositif, haptique, adaptable à divers cathéters, et utilisable en salle. Outre les difficultés de conception technique et de réalisation matérielle, un tel dispositif présentera cependant comme inconvénient de mesurer une force proximale, donc très éloignée de la zone distale du cathéter qui intéresse au premier plan le neuroradiologue interventionnel. Cela implique de savoir modéliser toute la zone intermédiaire (du point de ponction à l'aîne, jusqu'aux artères du cerveau), sur laquelle nous n'avons en général pas de données. Une alternative est d'équiper le cathéter de capteurs. La piste la plus prometteuse est celle de capteurs électromagnétiques insérés dans la paroi du cathéter. Les avancées sont cependant lentes dans ce domaine et il est peu probable que la miniaturisation soit suffisamment avancée à moyen terme pour permettre d'équiper des micro-cathéters. Enfin, l'approche la plus active aujourd'hui concerne les cathéters actifs, voire robotisés. La forme de ces cathéters peut être définie par un ensemble d'actuateurs mécaniques ou physiques localisés dans sa paroi. Le

lien entre partie proximale et partie distale serait ainsi plus aisé à maîtriser.

En second lieu, détecter signifie mesurer l'erreur entre la forme simulée et la forme visualisée. Dans le cas de la neuroradiologie interventionnelle, nous avons vu dans le Chapitre 3, Section 3.5 qu'un corpus bibliographique assez conséquent était à notre disposition pour détecter les micro-outils, souvent par le calcul d'une carte de caractéristiques, voire de probabilité de présence, avec une application au suivi de l'outil pour certains travaux. L'enjeu ici nous semble donc plutôt être de pouvoir prendre en compte une diversité de mesures d'erreur ou plus généralement de dissimilarité entre les formes simulées et visualisées, de telle manière qu'un lien direct peut être établi avec les paramètres de la simulation. On se rapproche ici des notions de matrice d'interaction en asservissement visuel ou d'innovation en filtrage bayésien. Ce point est donc plutôt en lien avec l'item «Adapter» ci-dessous. Des avancées dans ce domaine permettront cependant une meilleure évaluation des systèmes de simulation puisque l'interprétation des études de sensibilité en sera facilitée. Un point cependant sur le détecteur. Il nous semble important qu'il soit hybride et qu'il puisse prendre en compte à la fois la forme du cathéter et des marqueurs présents sur celui-ci (repères radio-opaques et/ou extrémité) afin d'être robuste au problème d'ouverture.

5.2.2 Prédire

Mieux prédire c'est mieux simuler.

La première piste de recherches vise à fixer les paramètres des modèles physiques, qui peuvent être mécaniques (module d'Young, coefficient de Poisson, etc...), mais aussi avoir une composante géométrique, par exemple le nombre de poutres pour un micro-guide, ou la forme au repos d'un micro-outil [26], ou encore les axes de rotation d'un modèle articulé des côtes [16]. Les deux citations en exemple indiquent deux voies. L'une consiste à optimiser une fonction de coût a posteriori d'une expérimentation. L'autre fait de même mais en temps réel et peut donc exploiter des données in-vivo. La difficulté commune à ces deux approches est le manque d'assurance que le mouvement et les changements de forme utilisés pour contraindre l'inversion de paramètres sont suffisamment discriminants pour restreindre l'espace des solutions à une seule paramétrisation. L'approche que nous souhaitons explorer est celle de l'assimilation de données dynamiques avec l'objectif d'informer et de guider l'expérimentateur, en exploitant les travaux sur les plans d'expérience séquentiels. La mise en œuvre se ferait ex-vivo, et concernerait l'outil déformable, avec l'objectif de mettre au point un banc de calibration léger et non destructif.

La deuxième piste de recherche s'intéresse à l'amélioration des modèles. Dans notre cas, il s'agit du modèle géométrique de l'organe. Nous avons déjà évoqué dans le Chapitre 4, Section 4.5 un certain nombre de perspectives d'amélioration du modèle rigide des vaisseaux sanguins. Des perspectives à plus long terme visent à l'élaboration d'un modèle déformable de ces vaisseaux. Tout d'abord, une déformation locale, liée à l'interaction avec l'outil, introduirait un peu de mollesse aux vaisseaux sanguins. Cette petite déformation est simplement réaliste, sans être nécessairement physique, et peut émaner d'un modèle de contact mou utilisé en animation graphique. Elle permettrait de déterminer non plus un point de contact, mais une zone de contact avec le micro-outil. La prise en compte de cette zone dans la résolution des contacts n'est pas triviale. Elle améliorerait cependant grandement le réalisme de la simulation en introduisant des contraintes latérales qui stabiliseraient la forme du micro-outil, et annuleraient les mouvements de pivot observés aujourd'hui autour du point de contact. Ensuite, un travail avec Pierre-Frédéric Villard de l'équipe Magrit pourrait adresser les déformations dues aux mouvements naturels, notamment la respiration. Nos travaux préliminaires [16] ouvrent la possibilité d'une approche guidée par l'image. Les difficultés à court terme consistent à trouver des partenaires médicaux, dont les données et le contexte clinique sont essentiels au soutien d'une telle recherche. Enfin, une perspective à plus long terme viserait à appréhender les grandes déformations, telles que celles qu'on peut observer en cathétérisme abdominal par exemple. Nous serons plus réservés sur le réalisme de cet objectif aujourd'hui car la modélisation physique de ce phénomène est extrêmement complexe. La mobilité des vaisseaux sanguins dépend en effet de divers facteurs, intrinsèques, avec des paramètres mécaniques spécifiques au patient et qui peuvent de plus varier

dans l'espace, par exemple au niveau de la pathologie, et extrinsèques, les vaisseaux sanguins étant souvent accolés à des organes dont la déformation va guider celle des vaisseaux.

5.2.3 Adapter

Adapter passe par la fusion des images avec le modèle physique simulé.

La première approche que nous suivons dans le cadre de la thèse de Raffaella Trivisonne applique à la simulation des contraintes issues des images. En pratique, un dispositif très pragmatique est utilisé où des marqueurs radio-opaques sont placés le long du cathéter, puis détectés et suivis par des algorithmes classiques de vision [12]. L'objectif premier est analytique et vise à mieux comprendre dans quelle mesure la simulation est sensible aux divers paramètres et contraintes, notamment de collision, et comment intégrer une contrainte projective dans le schéma numérique de résolution (méthode du lagrangien augmenté). Cet étape est préparatoire à l'utilisation raisonnée des approches de fusion bayésienne.

Nous souhaitons développer une approche bayésienne dans la suite de la thèse de Raffaella Trivisonne et à plus long terme. En effet, le cas s'y prête bien en théorie puisqu'on peut considérer que l'équation d'évolution s'implante dans l'exécution d'un pas de simulation (ou plusieurs, en fonction de la fréquence d'acquisition des données), et que l'équation d'observation fournit la mesure d'erreur évoquée dans la Section «Détecter» ci-dessus. La définition du vecteur d'état y est une question centrale. Bien sûr sa formulation influera sur le niveau de non-linéarité à gérer pour le calcul de l'innovation, mais surtout cette question nous met face à deux difficultés : préserver la cohérence du système mécanique, et en limiter la taille pour conserver des temps de calcul interactifs.

De manière classique en vision par ordinateur, le vecteur d'état contient les positions de points liés à l'objet, ce qui est en lien direct avec l'image qui permet d'observer, au moins partiellement, ces positions. Or le pilotage d'un système mécanique est plus naturel et stable en considérant les vitesses. Que ce soit dans un filtre de Kalman étendu (EKF), un filtre de Kalman sans parfum (UKF pour *Unscented Kalman Filter*) ou un filtre particulaire, les calculs reposent sur la génération d'instances du vecteur d'état : pour le calcul de la jacobienne avec l'EKF, la définition des sigma points dans l'UKF ou la génération des particules dans le filtre particulaire. Générer des instances des positions implique donc des vitesses infinies incompatibles avec les équations de la mécanique. De plus, le nombre d'instances à générer, et donc la vitesse d'exécution du filtre, dépend de la taille du vecteur d'état. Il convient donc de la limiter, ce qui peut se faire par une analyse de sensibilité ou par projection sur une base réduite (par exemple issue d'une Décomposition Propre Généralisée). Une autre piste est d'explorer les filtres de rang réduit. Cependant, il est probable qu'un filtre de type particulaire soit nécessaire pour gérer les cas où la distribution de probabilité est sans doute bi- (voire multi-) modale. La navigation d'un micro-cathéter à l'embranchement d'une bifurcation ou la transition d'un contact coller-glisser en sont deux exemples d'occurrence. Il n'existe pas de forme réduite de ces filtres à notre connaissance, mais la «malédiction de la dimensionnalité» a été abordée par un certain nombre de travaux sur lesquels nous pourrions nous baser. Dans le cadre du stage de Master de Maxime Malgras [Mal15], nous avons par exemple exploré le recours à une projection sur un chaos polynomial de la fonction implantée par un pas de simulation. Malheureusement, dans le cas où le chaos doit être appris à chaque pas de temps (chaos polynomial arbitraire), l'intérêt de cette approche n'apparaît pas pour ce qui est du temps de calcul. Dans le cadre de la neuroradiologie interventionnelle, où la dimension du vecteur pour décrire l'état du micro-outil semble plus accessible que pour un organe, notre banc de test et de calibration évoqué dans la section précédente nous aidera aussi à avancer sur cette question d'une paramétrisation minimale.

Bibliographie générale

- [ABK⁺96] J. H. Anderson, W. Brody, C. J. Kriz, Y. P. Wang, R. Raghavan, and R. Viswanathan. Da Vinci : A vascular catheterization and interventional radiology-based training and patient pretreatment planning simulator. *Journal of Vascular and Interventional Radiology*, 7(1) :373, January-February 1996.
- [ACC⁺02] J. H. Anderson, C. K. Chui, Y. Cai, Y. Wang, Z. Li, X. Ma, W. Nowinski, M. Solaiyappan, K. Murphy, P. Gailloud, and A. Venbrux. Virtual reality training in interventional radiology : The Johns Hopkins and Kent Ridge digital laboratory experience. *Seminars in Interventional Radiology*, 19(2) :179–185, 2002.
- [AdPP06] G. Aloisio, L. T. de Paolis, and L. Provenzano. A training simulator for the angioplasty intervention with a web portal for the virtual environment searching. In *Proceedings of the 5th WSEAS International Conference on Signal processing, Robotics and Automation*, pages 135–140, February 2006.
- [Anx03] René Anxionnat. *Méthodes et outils pour le détournement des malformations artérioveineuses cérébrales dans un contexte multimodalité*. PhD thesis, Ecole doctorale IAEM - Informatique, Automatique, Electronique - Electrotechnique, Mathématiques (Nancy), 2003.
- [APC⁺95] R.A. Alberico, M. Patel, S. Casey, B. Jacobs, W. Maguire, and R. Decker. Evaluation of the circle of Willis with three-dimensional CT angiography in patients with suspected intracranial aneurysms. *American Journal of Neuroradiology*, 16 :1571–1578, 1995.
- [AR98] J. H. Anderson and R. Raghavan. A vascular catheterization simulator for training and treatment planning. *Journal of Digital Imaging*, 11(3) :120–123, 1998.
- [Aro09] Michaël Aron. *Acquisition et modélisation de données articulatoires dans un contexte multimodal*. PhD thesis, Ecole doctorale IAEM - Informatique, Automatique, Electronique - Electrotechnique, Mathématiques (Nancy), 2009.
- [ATS95] N. Anzalone, F. Triulzi, and G. Scotti. Acute subarachnoid hemorrhage : 3D time-of flight MR angiography versus intra-arterial digital angiography. *Neuroradiology*, 37 :257–261, 1995.
- [BBH03] M.Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8) :993–1008, August 2003.
- [BBR⁺02] P. Badin, G. Bailly, L. Revéret, M. Baciú, C. Segerbarth, and C. Savariaux. Three-dimensional articulatory modeling of the tongue, lips and face, based on MRI and video images. *Journal of Phonetics*, 30(3) :533–553, 2002.
- [BDD09] M. Brückner, F. Deinzer, and J. Denzler. Temporal estimation of the 3D guide-wire position using 2D X-ray images. In *Proceedings of the 12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'2009)*, volume 5761 of *Lecture Notes in Computer Science*, pages 386–393, 2009.
- [BEB⁺07] P. Badin, F. Elisei, G. Bailly, C. Savariaux, A. Serrurier, and Y. Tarabalka. Têtes parlantes audiovisuelles virtuelles : données et modèles articulatoires - applications. *Revue de Laryngologie-Otologie-Rhinologie*, 128(5) :289–295, 2007.

- [BEG03] J. Beskow, O. Engwall, and B. Granström. Resynthesis of facial and intraoral motion from simultaneous measurements. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS'03)*, pages 431–434, 2003.
- [BF06] C. Beder and W. Förstner. Direct solutions for computing cylinders from minimal sets of points. In *9th European Conference on Computer Vision (ECCV'06)*, volume 3954 of *LNCS*, pages 135–146, May 2006.
- [BGGN91] T. Baer, J. C. Gore, L. C. Gracco, and P. W. Nye. Analysis of vocal tract shape and dimensions using magnetic resonance imaging : Vowels. *Journal of the Acoustical Society of America*, 90 :799–828, 1991.
- [BGK⁺08] F. Bender, M. Groher, A. Khamene, W. Wein, T. H. Heibel, and N. Navab. 3D dynamic roadmapping for abdominal catheterizations. In D. Metaxas, L. Axel, F. Fichtinger, and G. Székely, editors, *Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI'2008)*, volume 5242 of *Lecture Notes in Computer Science*, pages 668–675, 2008.
- [BGP⁺06] L. Brzozowski, A. Ganguly, M. Pop, Z. Wen, R. Bennet, R. Fahrig, and J. A. Rowlands. Compatibility of interventional X-ray and magnetic resonance imaging : Feasibility of a closed bore XMR (CBXMR) system. *Medical Physics*, 33 :3033–3045, 2006.
- [BH11] L. J. Brattain and R. D. Howe. Real-time 4D ultrasound mosaicing and visualization. In *Proceedings of the 14th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'11)*, volume 6891 of *Lecture Notes in Computer Science*, pages 105–112, 2011.
- [BKN⁺16] J. Banerjee, C. Klink, W. J. Niessen, A. Moelker, and T. van Walsum. 4D ultrasound tracking of liver and its verification for TIPS guidance. *IEEE Transactions on Medical Imaging*, 35(1) :52–62, 2016.
- [Bli82] J. F. Blinn. A generalization of algebraic surface drawing. *ACM Transactions on Graphics*, 1(3) :235–256, July 1982.
- [BLN⁺13] M. Berger, J. A. Levine, L. G. Nonato, G. Taubin, and C. T. Silva. A benchmark for surface reconstruction. *ACM Transactions on Graphics*, 32(2) :20 :1–17, 2013.
- [BM92] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) :239–256, 1992.
- [BmH01] R. Berthilsson, K. Åström, and A. Heyden. Reconstruction of general curves, using factorization and bundle adjustment. *International Journal of Computer Vision*, 41(3) :171–182, February/March 2001.
- [BMP⁺99] H.-J. Bender, R. Männer, C. Poliwoda, S. Roth, and M. Walz. Reconstruction of 3D catheter paths from 2D X-ray projection. In C. Taylor and A. Colchester, editors, *Proceedings of the 2nd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'1999)*, volume 1679 of *Lecture Notes in Computer Science*, pages 981–989. Springer-Verlag, Berlin, Heidelberg, 1999.
- [BRPP15] A. Bijar, P.Y. Rohan, P. Perrier, and Y. Payan. Atlas-based automatic generation of subject-specific finite element tongue meshes. *Annals of Biomedical Engineering*, 44(1) :16–34, 2015.
- [BTB⁺05] T. Bressmann, P. Thind, C.M. Bollig, C. Uy, R.W. Gilbert, and J.C. Irish. Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry : Data from twelve normal speakers and a partial glossectomee. *Clinical Linguistics and Phonetics*, 19 :573–588, 2005.
- [BTG95] E. Bittar, N. Tsingos, and M.-P. Gascuel. Automatic reconstruction of unstructured 3D data : Combining a medial axis and implicit surfaces. *Computer Graphics Forum*, 14(3) :457–468, 1995.

- [BvdKvW⁺03] S.A.M. Baert, E.B. van de Kraats, T. van Walsum, M. A. Viergever, and W.J. Niessen. 3D guide wire reconstruction from biplane image sequence for integrated display in 3D vasculature. *IEEE Transactions on Medical Imaging*, 22(10) :612–623, 2003.
- [BVG09] V. Bismuth, L. Vancamberg, and S. Gorges. A comparison of line enhancement techniques : applications to guide-wire detection and respiratory motion tracking. In *Proceedings of SPIE Medical Imaging 2009 : Visualization, Image-Guided Procedures, and Modeling*, page 72591M, 2009.
- [BVMA04] C. Blondel, R. Vaillant, G. Malandain, and N. Ayache. 3D tomographic reconstruction of coronary arteries using a precomputed 4D motion field. *Physics in Medicine and Biology*, 49(11) :2197–2208, 2004.
- [CCF⁺03] F. Colombo, C. Cavedon, P. Francescon, L. Casentini, U. Fornezza, L. Castellani, F. Causin, and S. Perini. Three-dimensional angiography for radiosurgical treatment planning for arteriovenous malformations. *Journal of Neurosurgery*, 98(3) :536–543, 2003.
- [CCY⁺03] Y. Cai, C. K. Chui, X. Ye, Y. Wang, and J. H. Anderson. VR simulated training for less invasive vascular intervention. *Computers & Graphics*, 27(2) :215–221, April 2003.
- [CDL⁺05] S. Cotin, C. Duriez, J. Lenoir, P. Neumann, and S. Dawson. New approaches to catheter navigation for interventional radiology simulation. In James S. Duncan and Guido Gerig, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'2005)*, volume 3750 of *Lecture Notes in Computer Science*, pages 534–542. Springer Berlin Heidelberg, 2005.
- [CDM⁺00] S. Cotin, S. L. Dawson, D. Meglan, D. W. Shaffer, M. A. Ferrell, R. S. Bardsley, F. M. Morgan, T. Nagano, J. Nikom, P. Sherman, M. T. Waltermann, and J. Wendlandt. ICTS, an interventional cardiology training system. *Studies in Health Technology and Informatics*, 70 :59–65, 2000.
- [CG01] T. Chaperon and F. Goulette. Extracting cylinders in full 3D data using a random sampling method and the Gaussian image. In *Vision Modeling and Visualization Conference (VMW'01)*, pages 35 – 42, 2001.
- [CHDO07] J. F. Carrillo, M. H. Hoyos, E. E. Dávila, and M. Orkisz. Recursive tracking of vascular tree axes in 3D medical images. *International Journal of Computer Assisted Radiology and Surgery*, 1(6) :331–339, 2007.
- [CJG⁺14] J.-F. Cai, W. Jia, H. Gao, S. B. Jiang, Z. Shen, and H. Zhao. Cine cone beam CT reconstruction using low-rank matrix factorization : Algorithm and a proof-of-principle study. *IEEE Transactions on Medical Imaging*, 33(8) :1581–1591, 2014.
- [CJL⁺99] T.S. Chung, J.Y. Joo, S.K. Lee, D. Chien, and G. Laub. Evaluation of intracranial aneurysms with high-resolution MR angiography using a section-interpolation technique : Correlation with digital subtraction angiography. *American Journal of Neuroradiology*, 20 :229–235, 1999.
- [CJT⁺00] H. J. Cloft, G. J. Joseph, F. C. Tong, J. H. Goldstein, and J. E. Dion. Use of three-dimensional Guglielmi detachable coils in the treatment of wide-necked cerebral aneurysms. *American Journal of Neuroradiology*, 21(7) :1312–1314, 2000.
- [CK41] T. Chiba and M. Kajiyama. *The Vowel : Its Nature and Structure*. Tokyo-Kaseikan, 1941.
- [CK01] V. Chalana and Y. Kim. A methodology for evaluation of boundary detection algorithms on medical images. *IEEE Transactions on Medical Imaging*, 16(5) :642–652, 2001.

- [CKP95] W. J. Christmas, J. Kittler, and M. Petrou. Structural matching in computer vision using probabilistic relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8) :749–764, August 1995.
- [CMD⁺95] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multimodality medical image registration using information theory. In Y. Bizais and C. Barillot, editors, *Proceedings of the 14th International Conference on Information Processing in Medical Imaging (IPMI'95)*, pages 263–274, 1995.
- [CMM⁺16] M. E. Currie, A. J. McLeod, J. T. Moore, M. W. A. Chu, R. Patel, B. Kiaii, and T. M. Peters. Augmented reality system for ultrasound guidance of transcatheter aortic valve implantation. *Innovations : Technology & Techniques in Cardiothoracic & Vascular Surgery*, 11(1) :31–39, 2016.
- [CNMR02] C. Cañero, E. Nofrerías, J. Mauri, and P. Radeva. Modelling the acquisition geometry of a C-arm angiography system for 3D reconstruction. In *Proceedings of the 5th Catalanian Conference on Artificial Intelligence (CCIA'2002)*, volume 2504 of *Lecture Notes in Computer Science*, pages 322–335, Castellón, Spain, 2002.
- [CP10] K. Cleary and T. M. Peters. Image-guided interventions : Technology review and clinical applications. *Annual Review of Biomedical Engineering*, 12 :119–142, 2010.
- [CS09] V. Chari and P. Sturm. Multiple view geometry of the refractive plane. In *Proceedings of the 20th British Machine Vision Conference (BMVC'2009)*, pages 1–11, 2009.
- [CTCG95] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1) :38–59, January 1995.
- [CTEA09] T.K. Chen, A.D. Thurston, R.E. Ellis, and P. Abolmaesumi. A real-time freehand ultrasound calibration system with automatic accuracy feedback and control. *Ultrasound in Medicine and Biology*, 35(1) :79–93, January 2009.
- [CTM08] S. Curtis, R. Tamstorf, and D. Manocha. Fast collision detection for deformable models using representative-triangles. In *Proceedings of the 2008 Symposium on Interactive 3D Graphics and Games (I3D'2008)*, pages 61–69, New York, NY, USA, 2008. ACM.
- [CV01] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2) :266–277, February 2001.
- [CVMR02] C. Cañero, F. Vilariño, J. Mauri, and P. Radeva. Predictive (un)distortion model and 3-D reconstruction by biplane snakes. *IEEE Transactions on Medical Imaging*, 21(9) :1188–1201, September 2002.
- [CWC⁺98] C. Cognard, A. Weill, L. Castaings, A. Rey, and J. Moret. Intracranial berry aneurysms : Angiographic results after endovascular treatment. *Radiology*, 206 :499–510, 1998.
- [DA00] J. C. Duncan and N. Ayache. Medical image analysis : Progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1) :85–106, January 2000.
- [DAK04] C. Duriez, C. Andriot, and A. Kheddar. Signorini’s contact model for deformable objects in haptic simulations. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'2004)*, volume 4, pages 3232–3237, 2004.
- [Dar04] F. Darema. Dynamic data driven applications systems : A new paradigm for application simulations and measurements. In *Proceedings of the International*

- Conference on Computational Science (ICCS'2004)*, volume 3038 of *Lecture Notes in Computer Science*, pages 662–669, 2004.
- [Daw06] S. Dawson. Procedural simulation : A primer. *Radiology*, 241(1) :17–25, 2006.
- [DC02] T. Deschamps and L.D. Cohen. Fast extraction of tubular and tree 3D surfaces with front propagation methods. In *Proceedings of the 16th International Conference on Pattern Recognition (ICPR'2002)*, volume 1, pages 731–734 vol.1, 2002.
- [DCLN06] C. Duriez, S. Cotin, J. Lenoir, and P. F. Neumann. New approaches to catheter navigation for interventional radiology simulation. *Computer Aided Surgery*, 11(6) :300–308, 2006.
- [Del17] Charlotte Delmas. *Reconstruction 3D de micro-outils curvilignes en neuroradiologie interventionnelle*. PhD thesis, Ecole doctorale IAEM - Informatique, Automatique, Electronique - Electrotechnique, Mathématiques (Nancy), 2017.
- [Dem04] Sébastien Demange. Minimisation d'énergie par graph-cuts appliquée à la segmentation d'images et à la stéréoscopie. Master's thesis, École Doctorale IEAM Lorraine, 2004.
- [DK98] S. L. Dawson and J. A. Kaufman. The imperative for medical simulation. *Proceedings of the IEEE*, 86(3) :479–483, march 1998.
- [DLC07] J. Dequidt, J. Lenoir, and S. Cotin. Interactive contacts resolution using smooth surface representation. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI'2007)*, volume 4792 of *Lecture Notes in Computer Science*, pages 850–857. Springer-Verlag, Berlin, Heidelberg, 2007.
- [DNH16] M. M. Dalvand, S. Nahavandi, and R. D. Howe. Fast vision-based catheter 3D reconstruction. *Physics in Medicine and Biology*, 61 :5128–5148, 2016.
- [DRG94] A.C.M. Dumay, J.H.C. Reiber, and J.J. Gerbrands. Determination of optimal angiographic viewing angles : Basic principles and evaluation study. *IEEE Transactions on Medical Imaging*, 13(1) :13–24, March 1994.
- [DTC⁺02] R.H. Davies, C.J. Twining, T.F. Cootes, J.C. Waterton, and C.J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5) :525–537, 2002.
- [DvOG04] D. Dekkers, K. van Overveld, and R. Golsteijn. Combining CSG modeling with soft blending using Lipschitz-based implicit surfaces. *The Visual Computer*, 20(6) :380–391, 2004.
- [EJH⁺04] P.J. Edwards, L.G. Johnson, D.J. Hawkes, M.R. Fenlon, A.J. Strong, and M.J. Gleeson. Clinical experience and perception in stereo augmented reality surgical navigation. In *Proceedings of the Conference on Medical Imaging Augmented Reality (MIAR'2004)*, pages 369–376, 2004.
- [Eng04] O. Engwall. From real-time MRI to 3D tongue movements. In Soon Hyob Kim and Dae Hee Youn, editors, *Proceedings of the 8th International Conference on Spoken Language Processing (Interspeech'2004)*, pages 1109–1112, Jeju Island, Corée du Sud, Octobre 2004.
- [Eng08] O. Engwall. Can audio-visual instructions help learners improve their articulation ? - an ultrasound study of short term changes. In *Proceedings of the 9th Annual Conference of the International Speech Communication Association (Interspeech'2008)*, pages 2631–2634, 2008.
- [Fan60] G. Fant. *Acoustic Theory of Speech Production*. The Hague : Mouton & Co., 1960.
- [Fau93] O.D. Faugeras. *Three-dimensional computer vision : A geometric viewpoint*. MIT Press, 1993.

- [FB81] M. A. Fischler and R. C. Bolles. Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24 :381–395, 1981.
- [FBW⁺16] P. Favollita, A. Brand, L. Wang, E. Euler, P. Thaller, N. Navab, and S. Weidert. An augmented reality C-arm for intraoperative assessment of the mechanical axis : a preclinical study. *International Journal of Computer Assisted Radiology and Surgery*, 11 :2111–2117, 2016.
- [FDCR01] N. Flasque, M. Desvignes, J.-M. Constans, and M. Revenu. Acquisition, segmentation and tracking of the cerebral vascular tree on 3D magnetic resonance angiography images. *Medical Image Analysis*, 5(3) :173–183, 2001.
- [FDD⁺12] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterlik, and S. Cotin. SOFA : A multi-model framework for interactive physical simulation. In Yohan Payan, editor, *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, volume 11 of *Studies in Mechanobiology, Tissue Engineering and Biomaterials*, pages 283–321. Springer, 2012.
- [FDM⁺04] G. Fichtinger, A. Deguet, K. Masamune, G. Fischer, E. Balogh, H. Mathieu, R.H. Taylor, L.M. Fayad, and S.J. Zinreich. Needle insertion in CT scanner with image overlay - cadaver studies. In *Proceedings of the 7th International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI'2004)*, volume 3217 of *Lecture Notes in Computer Science*, pages 795–803. Springer-Verlag, 2004.
- [FFLH97] R. Fahrig, A.J. Fox, S. Lownie, and D.W. Holdsworth. Use of C-arm system to generate true three-dimensional computed rotational angiograms : Preliminary in-vitro and in-vivo results. *American Journal of Neuroradiology*, 18 :1507–1514, 1997.
- [FHKP10] O. Friman, M. Hindennach, C. Kühnel, and H.-O. Peitgen. Multiple hypothesis template tracking of small 3D vessel structures. *MedIA*, 14(2) :160 – 171, 2010.
- [Fit09] J.M. Fitzpatrick. Fiducial registration error and target registration error are uncorrelated. In M.I. Miga and K.H. Wong, editors, *Proceedings of SPIE Medical Imaging 2009 : Visualization, Image-Guided Procedures, and Modeling*, volume 7261, page 726102, 2009.
- [FK10] R. Fabbri and B. Kimia. 3D curve sketch : Flexible curve-based stereo reconstruction and calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2010)*, pages 1538–1545, San Francisco, CA, USA, 2010.
- [FKL⁺99] J.C. Flickinger, D. Kondziolka, L.D. Lunsford, B.E. Pollock, M. Yamamoto, D.A. Gorman, P.J. Schomberg, P. Sneed, D. Larson, V. Smith, M.W. McDermott, L. Miyawaki, J. Chilton, R.A. Morantz, B. Young, H. Jokura, and R. Liscak. A multi-institutional analysis of complication outcomes after arteriovenous malformation radiosurgery. *International Journal of Radiation Oncology, Biology, Physics*, 44(1) :67–74, April 1999.
- [FWB04] P. Felkel, R. Wegenkittl, and K. Buhler. Surface models of tube trees. In *Proceedings of the Computer Graphics International (CGI'2004)*, pages 70–77, Washington, DC, USA, 2004. IEEE Computer Society.
- [FWMJ98] J.M. Fitzpatrick, J.B. West, and C. R. Maurer Jr. Predicting error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, 17(5) :694–702, 1998.
- [Gas93] M.-D. Gascuel. An implicit formulation for precise contact modeling between flexible solids. In *Proceedings of SIGGRAPH 93*, Annual Conference Series, pages 313–320, August 1993.

- [GBC⁺13] O. Gourmel, L. Barthe, M.-P. Cani, B. Wyvill, A. Bernhardt, M. Paulin, and H. Grasberger. A gradient-based implicit blend. *ACM Transactions on Graphics*, 32(2), 2013. Article No. 12.
- [GBK⁺09] M. Groher, F. Bender, A. Khamene, W. Wein, T. H. Heibel, and N. Navab. 3D guide wire navigation from single plane fluoroscopic images in abdominal catheterizations. In *Proceedings of the 12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'2009)*, volume 5761 of *Lecture Notes in Computer Science*, 2009.
- [GCJ⁺12] D. Gould, N. Chalmers, S. J. Johnson, C. Kilkenney, M. D. White, B. Bech, L. Lonn, and F. Bello. Simulation : Moving from technology challenge to human factors success. *CardioVascular and Interventional Radiology*, 35(3) :445–453, June 2012.
- [Gla09] P. Glanc. Modélisation de vaisseaux sanguins par surfaces implicites. Master’s thesis, TELECOM ParisTech, 2009.
- [GMM⁺98] P. Gallina, L. Merienne, J.F. Meder, M. Schlienger, D. Lefkopoulos, and J.J. Merland. Failure in radiosurgery treatment of cerebral arteriovenous malformations. *Neurosurgery*, 42(5) :996–1004, 1998.
- [Gol05] R. Goldman. Curvature formulas for implicit curves and surfaces. *Computer Aided Geometric Design*, 22(7) :632–658, October 2005.
- [Gor07] Sébastien Gorges. *Vers un système de navigation 3D en neuroradiologie interventionnelle*. PhD thesis, Ecole doctorale IAEM - Informatique, Automatique, Electronique - Electrotechnique, Mathématiques (Nancy), 2007.
- [Gou10] D. Gould. Using simulation for interventional radiology training. *The British Journal of Radiology*, 83 :546–553, 2010.
- [GW99] Y. Gdalyahu and D. Weinshall. Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(12) :1312–1328, 1999.
- [Hao15] N. Haouchine. *Image-guided simulation for augmented reality during hepatic surgery*. PhD thesis, Doctoral School of Engineering Science, Lille, France, 2015.
- [HBJ⁺12] M. Hoffman, A. Brost, C. Jakob, F. Bourier, M. Koch, K. Kurzidim, J. Hornegger, and N. Strobel. Semi-automatic catheter reconstruction from two views. In *Proceedings of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'2012)*, volume 7511 of *Lecture Notes in Computer Science*, pages 584–591, 2012.
- [HCG⁺98] J.F. Heautot, E. Chabert, Y. Gandon, S. Croci, R. Roméas, R. Campagnolo, B. Chereul, J.M. Scarabin, and M. Carsin. Analysis of cerebrovascular diseases by a new 3-dimensional computerised X-ray angiography system. *Neuroradiology*, 40 :203–209, 1998.
- [HLT12] Q. Hong, Q. Li, and J. Tian. Implicit reconstruction of vasculatures using bivariate piecewise algebraic splines. *IEEE Trans. Med. Imaging*, 31(3) :543–553, march 2012.
- [HM95] R. Horaud and O. Monga. *Vision par ordinateur, outils fondamentaux*. Éditions Hermès, 1995.
- [HMM⁺95] G. A. Higgins, D. A. Meglan, A. S. Millman, R. Horst, G. L. Merrill, and J. R. Merrill. Virtual reality surgery : Implementation of a coronary angioplasty training simulator. *Cardiovascular Surgery Surgical Technology International*, IV :379–383, 1995.
- [Hol08] M. Holden. A review of geometric transformations for nonrigid body registration. *IEEE Transactions on Medical Imaging*, 27(1) :111–128, 2008.

- [Hol15] K. G. Hollingsworth. Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction. *Physics in Medicine and Biology*, 60(21) :292–322, 2015.
- [HOPJ⁺06] M. H. Hoyos, P. Orlowski, E. Piatkowska-Janko, P. Bogorodzki, and M. Orkisz. Vascular centerline extraction in 3D MR angiograms for phase contrast MRI blood flow measurement. *International Journal of Computer Assisted Radiology and Surgery*, 1(1) :51–61, 2006.
- [HPSP01] H. K. Hahn, B. Preim, D. Selle, and H.-O. Peitgen. Visualization and interaction techniques for the exploration of vascular structures. In *Visualization (VIS'2001)*, pages 395–402, 2001.
- [HS89] R. Horaud and T.H. Skordas. Stereo correspondence through feature grouping and maximal cliques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(11) :1168–1180, November 1989.
- [HWtG94] D.J. Hoff, M.C. Wallace, K.G. terBrugge, and F. Gentili. Rotational angiography assessment of intracranial aneurysms. *American Journal of Neuroradiology*, 15 :1945–1948, 1994.
- [HZ00] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN : 0521623049, 2000.
- [IFV⁺15] P. W. Iltis, J. Frahm, D. Voit, A. A. Joseph, E. Schoonderwaldt, and E. Altenmüller. High-speed real-time magnetic resonance imaging of fast tongue movements in elite horn players. *Quantitative Imaging in Medicine and Surgery*, 5(3) :374–381, June 2015.
- [IKY97] M. Ida, Y. Kurisu, and M. Yamashita. MR angiography of ruptured aneurysms in acute subarachnoid hemorrhage. *American Journal of Neuroradiology*, 18 :1025–1032, 1997.
- [JEH02] L.G. Johnson, P.G. Edwards, and D.J. Hawkes. Surface transparency makes stereo overlays unpredictable : the implications for augmented reality. In *Proceedings of the Conference Medicine Meets Virtual Reality (MMVR'2002)*, pages 131–136, 2002.
- [KA03] F. Kahl and J. August. Multiview reconstruction of space curves. In *Proceedings of the 9th International Conference on Computer Vision (ICCV'2003)*, volume 2, pages 1017–1024, 2003.
- [KBH06] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Proceedings of the 4th Eurographics symposium on Geometry processing (SGP'2006)*, pages 61–70, Aire-la-Ville, Switzerland, 2006. Eurographics Association.
- [KBS06] W. F. Katz, S. V. Bharadwaj, and M. P. Stettler. Influences of electromagnetic articulography sensors on speech produced by healthy adults and individuals with aphasia and apraxia. *Journal of Speech, Language and Hearing Research*, 49 :645–659, 2006.
- [KGPS13] J. Kretschmer, C. Godenschwager, B. Preim, and M. Stamminger. Interactive patient-specific vascular modeling with sweep surfaces. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12) :2828–2837, Dec 2013.
- [KHI⁺07] S. Kockara, T. Halic, K. Iqbal, C. Bayrak, and R. Rowe. Collision detection : A survey. In *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, pages 4046–4051, 2007.
- [KNN09] Y. C. Kim, S. S. Narayanan, and K. S. Nayak. Accelerated 3D MRI of vocal tract shaping using compressed sensing and parallel imaging. In *Proceedings of the 34th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'2009)*, pages 389–392, Taipei, Taiwan, 2009.

- [KOJC12] M. Kersten-Oertel, P. Jannin, and D. L. Collins. DVV : A taxonomy for mixed reality visualization in image guided surgery. *IEEE Transactions on Visualization and Computer Graphics*, 18(2) :332–352, February 2012.
- [KS04] J.Y. Kaminski and A. Shashua. Multiple view geometry of general algebraic curves. *International Journal of Computer Vision*, 56(3) :195–2019, 2004.
- [KS05] A. Khamene and F. Sauer. A novel phantom-less spatial and temporal ultrasound calibration method. In *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'05)*, volume 3750 of *LNCS*, pages 65–72, 2005.
- [KTM⁺96] Y. Korogi, M. Takahashi, N. Mabuchi, T. Nakagawa, S. Fujiwara, Y. Horikawa, H. Miki, T. O'Uchi, H. Shiga, Y. Shiokawa, T. Watabe, and M. Furuse. Intracranial aneurysms : Diagnostic accuracy of MR angiography with evaluation of maximum intensity projection and source images. *Radiology*, 199(1) :199–207, 1996.
- [KZ01] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *Proceedings of the 8th International Conference on Computer Vision (ICCV'01)*, pages 508–515, 2001.
- [KZ04] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph-cuts ? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2) :147–159, February 2004.
- [LAB⁺09] D. Lesage, E.D. Angelini, I. Bloch, et al. A review of 3D vessel lumen segmentation techniques : Models, features and extraction schemes. *Medical Image Analysis*, 13(6) :819–845, 2009.
- [Lau95] A. Laurentini. How far 3D shapes can be understood from 2D silhouettes ? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2) :188–195, 1995.
- [Lau96] Laurent Launay. *Localisation et reconstruction 3D à partir d'angiographies stéréotaxiques*. PhD thesis, Ecole doctorale IAEM - Informatique, Automatique, Electronique - Electrotechnique, Mathématiques (Nancy), 1996.
- [LBV⁺16] M. Labrunie, P. Badin, D. Voit, A.A. Joseph, L. Lamalle, C. Vilain, L.-J. Boë, and J. Frahm. Tracking contours of orofacial articulators from real-time MRI of speech. In *Proceedings of the 17th Annual Conference of the International Speech Communication Association (Interspeech'2016)*, pages 470–474, 2016.
- [LCSK⁺04] A. La Cruz, M. Straka, A. Kochl, M. Sramek, E. Groller, and D. Fleischmann. Non-linear model fitting to parameterize diseased blood vessels. In *Proceedings of the conference on Visualization (VIS'2004)*, pages 393–400, Washington, DC, USA, 2004. IEEE Computer Society.
- [LDC⁺13] C. A. Linte, K. P. Davenport, K. Cleary, C. Peters, K. G. Vosburgh, N. Navab, P.G. Edwards, P. Jannin, T. M. Peters, D. R. Holmes III, and R. A. Robb. On mixed reality environments for minimally invasive therapy guidance : Systems architecture, successes and challenges in their implementation from laboratory to clinic. *Computerized Medical Imaging and Graphics*, 37 :83–97, 2013.
- [Lee00] I.-K. Lee. Curve reconstruction from unorganized points. *Computer Aided Geometric Design*, 17(2) :161–177, 2000.
- [Lem10] V. Lempitsky. Surface extraction from binary volumes with higher-order smoothness. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2010)*, pages 1197 –1204, 2010.
- [LG98] M. Lin and S. Gottschalk. Collision detection between geometric models : A survey. In *Proceedings of IMA conference on mathematics of surfaces*, volume 1, pages 602–608, 1998.

- [LHG⁺09] V. Luboz, C. Hughes, D. Gould, N. John, and F. Bello. Real-time Seldinger technique simulation in complex vascular models. *International Journal of Computer Assisted Radiology and Surgery*, 4 :589–596, November 2009.
- [Lin03] Q. Lin. *Enhancement, detection and visualization of 3D volume data*. PhD thesis, Linköping University, 2003.
- [LKGD07] C. Li, C. Kao, J. Gore, and Z. Ding. Implicit active contours driven by local binary fitting energy. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR’2007)*, 2007.
- [LKTS⁺13] V. Luboz, J. Kyaw-Tun, S. Sen, R. Kneebone, R. Dickinson, R. Kitney, and F. Bello. Real-time stent and balloon simulation for stenosis treatment. *The Visual Computer*, pages 1–9, 2013.
- [LLSK10] R. Liao, D. Luc, Y. Sun, and K. Kirchberg. 3-D reconstruction of the coronary artery tree from multiple views of a rotational X-ray angiography. *International Journal of Cardiovascular Imaging*, 26(7) :733–749, October 2010.
- [LM99b] J.O. Lachaud and A. Montanvert. Deformable meshes with automated topology changes for coarse to fine 3D surface extraction. *Medical Image Analysis*, 3(2) :187–207, 1999.
- [LVD98] J.M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *Proceedings of the European Conference on Computer Vision (ECCV’1998)*, volume 1, pages 158–174, Freiburg, Germany, 1998.
- [LYTM06] C. Lauterbach, S.-E. Yoon, D. Tuft, and D. Manocha. RT-DEFORM : Interactive ray tracing of dynamic scenes using BVHs. In *Interactive Ray Tracing 2006, IEEE Symposium on*, pages 39–46, Sept 2006.
- [LZ06] G. Li and S. W. Zucker. Contextual inference in contour-based stereo correspondence. *International Journal of Computer Vision*, 69(1) :59–75, 2006.
- [LZZ⁺15] J. Liu, X. Zhang, X. Zhang, H. Zhao, Y. Gao, D. Thomas, D. A. Low, and H. Gao. 5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography. *Inverse Problems*, 31(11) :(21 pp), 2015.
- [Ma07] X. Ma. Latest development of an interventional radiology training simulation system : Neurocath. In V. G. Duffy, editor, *Proceedings of the International Conference on Digital Human Modeling (ICDHM’2007)*, volume 4561 of *Lecture Notes in Computer Science*, pages 684–693. Springer, Berlin, Heidelberg, 2007.
- [Mae79] S. Maeda. Un modèle articulatoire de la langue avec des composantes linéaires. In *Actes 10èmes Journées d’Etude sur la Parole*, pages 152–162, 1979.
- [Mal15] Maxime Malgras. Simulation guidée par l’image. Master’s thesis, École Doctorale IEAM Lorraine, 2015.
- [MBB⁺17] S. Matl, R. Brosig, M. Baust, N. Navab, and S. Demirci. Vascular image registration techniques : A living review. *Medical Image Analysis*, 35 :1–17, 2017.
- [MBE⁺06] S. Maeda, M. O. Berger, O. Engwall, Y. Laprie, P. Maragos, B. Potard, and J. Schoentgen. Technology inventory of audiovisual-to-articulatory inversion. Technical report, FET ASPI Report (EU Project no. 2005-021324), 2006.
- [MCK⁺13] J. T. Moore, M. W. A. Chu, B. Kiaii, D. Bainbridge, G. Guiraudon, C. Wedlake, M. Currie, M. Rajchl, R. V. Patel, and T. M. Peters. A navigation platform for guidance of beating heart transapical mitral valve repair. *IEEE Transactions on Biomedical Engineering*, 60(4) :1034–1040, 2013.
- [MdGD⁺10] P. Mullen, F. de Goes, M. Desbrun, D. Cohen-Steiner, and P. Alliez. Signing the unsigned : Robust surface reconstruction from raw pointsets. *Computer Graphics Forum*, 29(5) :1733–1741, 2010.

- [Meg96] D. Meglan. Making surgical simulation real. *Computer Graphics*, 30(4) :37–39, November 1996.
- [MGBL07] H. Martinsson, F. Gaspard, A. Bartoli, and J. M. Lavest. Energy-based reconstruction of 3D curves for quality control. In A.L. Yuille, S.C. Zhu, D. Cremers, and Y. Wang, editors, *Proceedings of the 6th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCV-PR'2007)*, volume 4679 of *Lecture Notes in Computer Science*, pages 414–428. Springer, Berlin, Heidelberg, 2007.
- [MHH⁺05] R. A. McLaughlin, J. Hipwell, D. J. Hawkes, J. A. Noble, J. V. Byrne, and T. C. Cox. A comparison of a similarity-based and a feature-based 2D-3D registration method for neurointerventional use. *IEEE Transactions on Medical Imaging*, 24(8) :1058–1066, 2005.
- [MHV⁺03] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1) :120–128, Janvier 2003.
- [MIH11] D. J. Mirota, M. Ishii, and G. D. Hager. Vision-based navigation in image-guided interventions. *Annual Review of Biomedical Engineering*, 13 :297–319, 2011.
- [MLLC05] L. Mercier, T. Langø, F. Lindseth, and L. D. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine and Biology*, 31(2) :143–165, February 2005.
- [MPLŠ15] U. Mitrović, F. Pernuš, B. Likar, and Ž. Špiclin. Simultaneous 3D-2D image registration and C-arm calibration : Application to endovascular image-guided interventions. *Medical Physics*, 42(11) :6433–6447, November 2015.
- [MTLP12] P. Markelj, D. Tomazčević, B. Likar, and F. Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, 16 :642–661, 2012.
- [MTLP13] U. Mitrović, D. Tomazčević, B. Likar, and F. Pernuš. 3D-2D registration of cerebral angiograms : A method and evaluation on clinical images. *IEEE Transactions on Medical Imaging*, 32(8) :1550–1563, August 2013.
- [Mur91] S. Muraki. Volumetric shape description of range data using blobby model. *Computer Graphics*, 25 :227–235, 1991.
- [MV98] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1) :1–16, 1998.
- [NAAR14] M. Najafi, N. Afsham, P. Abolmaesumi, and R. Rohling. A closed-form differential formulation for ultrasound spatial calibration : Multi-wedge phantom. *Ultrasound in Medicine and Biology*, 40(9) :2231–2243, September 2014.
- [NAAR15] M. Najafi, N. Afsham, P. Abolmaesumi, and R. Rohling. A closed-form differential formulation for ultrasound spatial calibration : Single wall phantom. *Ultrasound in Medicine and Biology*, 41(4) :1079–1094, April 2015.
- [NC01] W. L. Nowinski and C. K. Chui. Simulation of interventional neuroradiology procedures. In *Proceedings of the International Workshop on Medical Imaging and Augmented Reality (MIAR'2001)*, pages 87–94, Shatin, Hong Kong, June 2001.
- [NF15] I. Nurutdinova and A. W. Fitzgibbon. Towards pointless structure from motion : 3D reconstruction and camera parameters from general 3D curves. In *Proceedings of the International Conference on Computer Vision (ICCV'2015)*, pages 2363–2371, 2015.
- [NFF16] C. Navab, N. ad Hennersperger, B. Frisch, and B. Fürst. Personalized, relevance-based multimodal robotic imaging and augmented reality for computer assisted interventions. *Medical Image Analysis*, 33 :64–71, October 2016.

- [NHT10] N. Navab, S.M. Heining, and J. Traub. Camera augmented mobile C-arm (CAMC) : Calibration, accuracy study and clinical applications. *IEEE Transactions on Medical Imaging*, 29(7) :1412–1423, 2010.
- [NPSA04] S. Nicolau, X. Pennec, L. Soler, and N. Ayache. An accuracy certified augmented reality system for therapy guidance. In *Proceedings of the 8th European Conference on Computer Vision (ECCV'2004)*, volume 3023, pages 79–91, Prague, Czech Republic, May 2004. Springer.
- [OAA⁺12] Y. Otake, M. Armand, R. S. Armiger, M. D. Kutzer, E. Basafa, P. Kazanzides, and R. H. Taylor. Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery : Incorporation of fiducial-based C-arm tracking and GPU-acceleration. *IEEE Transactions on Medical Imaging*, 31(4) :948–962, April 2012.
- [OBA⁺03] Y. Ohtake, A. Belyaev, M. Alexa, G. Turk, and H.-P. Seidel. Multi-level partition of unity implicits. *ACM Transactions on Graphics*, 22(3) :463–470, 2003.
- [OP05] S. Oeltze and B. Preim. Visualization of vasculature with convolution surfaces : Method, validation and evaluation. *IEEE Transactions on Medical Imaging*, 24(4) :540–548, 2005.
- [PBC⁺06] G.P. Penney, D.C. Barratt, C.S.K. Chan, M. Slomczykowski, T.J. Carter, P.J. Edwards, and D.J. Hawkes. Cadaver validation of intensity-based ultrasound to CT registration. *Medical Image Analysis*, 10(3) :385–395, 2006.
- [PCS⁺92] J. S. Perkell, M. H. Cohen, M. A. Svirsky, M. L. Matthies, I. Garabieta, and M. T. T. Jackson. Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *Journal of Acoustical Society of America*, 92(6) :3078–3096, 1992.
- [PD17] R. Patel and R. Dennick. Simulation based teaching in interventional radiology training : Is it effective ? *Clinical Radiology*, 72(3) :266.e7–266.e14, March 2017.
- [PF15] X. Pennec and P. Fillard. Statistical computing on non-linear spaces for computational anatomy. In N. Paragios, J. Duncan, and N. Ayache, editors, *Handbook of Biomedical Imaging*, pages 147–168. Springer US, 2015.
- [PFJ05] P. Paul, O.F. Fleig, and P. Jannin. Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery : methods and performance evaluation. *IEEE Transactions on Medical Imaging*, 24(11) :1500–1511, 2005.
- [PFL⁺98] B.E. Pollock, J.C. Flickinger, L.D. Lunsford, A. Maitz, and D. Kondziolka. Factors associated with successful arteriovenous malformations radiosurgery. *Neurosurgery*, 42(6) :1244–1247, 1998.
- [PH02] M. Personnaz and R. Horaud. Camera calibration : Estimation, validation and software. Rapport de recherche RT-0258, Inria, 2002.
- [PLDH11] T. Petkovič, S. Lončarić, T. Devčić, and R. Homan. Non-iterative guidewire reconstruction from multiple projective views. In *Proceedings of the 7th Symposium on Image and Signal Processing and Analysis (ISPA '2011)*, pages 639–643, 2011.
- [PMV03] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images : A survey. *IEEE Transactions on Medical Imaging*, 22(8) :986–1004, Août 2003.
- [PO08] B. Preim and S. Oeltze. 3D visualization of vasculature : An overview. In Lars Linsen, Hans Hagen, Bernd Hamann, Gerald Farin, Hans-Christian Hege, David Hoffman, Christopher R. Johnson, Konrad Polthier, and Martin Rumpf, editors, *Visualization in Medicine and Life Sciences*, Mathematics and Visualization, pages 39–59. Springer Berlin Heidelberg, 2008.

- [PRBT17] L. Picard, G. Rodesch, S. Bracard, and A. Taylor. Recommendation of the WFITN regarding simulation in neurointerventional training. *Interventional Neuroradiology*, 2017.
- [PRGB98] R.W. Prager, R.N. Rohling, A.H. Gee, and L Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6) :855–869, March 1998.
- [PS90] P. Puget and T. H. Skordas. An optimal solution for mobile camera calibration. In *Proceedings of the European Conference on Computer Vision (ECCV’1990)*, pages 187–198, Antibes, France, 1990.
- [PT97] X. Pennec and J. P. Thirion. A framework for uncertainty and validation of 3-D registration methods based on points and frames. *International Journal of Computer Vision*, 25(3) :203–229, December 1997.
- [RBSH04] P. Rebholz, C. Bienek, D. Stsepankou, and J. Hesser. CathI – Training system for PTCA. a step closer to reality. In S. Cotin and D. Metaxas, editors, *Medical Simulation (ISMS’2004)*, volume 3078 of *Lecture Notes in Computer Science*, pages 249–255. Springer, Berlin, Heidelberg, 2004.
- [RF91] L. Robert and O.D. Faugeras. Curve-based stereo : figural continuity and curvature. In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR’1991)*, pages 57–62, Maui, Hawai, USA, 1991.
- [RHM⁺11] D. Ruijters, R. Homan, P. Mielekamp, P. van de Haar, and Babic D. Validation of 3D multimodality roadmapping in interventional neuroradiology. *Physics in Medicine and Biology*, 56 :5335–5354, August 2011.
- [RPTP93] A. Rougée, C. Picard, Y. Troussel, and C. Ponchut. Geometrical calibration for 3D X-ray imaging. In Y. Kim, editor, *Proceedings of SPIE Medical Imaging 1993*, volume 1897, pages 161–169. SPIE Press, June 1993.
- [RSB⁺05] K. S. Rhode, M. Sermesant, D. Brogan, S. Hegde, J. Hipwell, P. Lambiase, E. Rosenthal, C. Bucknall, S. A. Qureshi, J. S. Gill, R. Razavi, and D. L. G Hill. A system for real-time XMR guided cardiovascular intervention. *IEEE Transactions on Medical Imaging*, 24(11) :1428–1440, 2005.
- [RTMT10] S. Ramalingam, Y. Taguchi, T. K. Marks, and O. Tuzel. $p2\pi$: A minimal solution for registration of 3D points to 3D planes. In *Proceedings of the 11th European Conference on Computer Vision (ECCV’10)*, volume 6315 of *LNCS*, pages 436–449, 2010.
- [SBHA05] M. Söderman, D. Babic, R. Homan, and T. Andersson. 3D roadmap in neuroangiography : technique and clinical interest. *Neuroradiology*, 47(10) :735–740, 2005.
- [SBHN06] T. Sielhorst, C. Bichlmeier, S. Heining, and N. Navab. Depth perception a major issue in medical AR : Evaluation study by twenty surgeons. In R. Larsen, M. Nielsen, and J. Sporring, editors, *Proceedings of the 9th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI’2006)*, volume 4190 of *Lecture Notes in Computer Science*, pages 364–372, Copenhagen, Denmark, 2006.
- [SBS03] R. Sierra, M. Bajka, and G. Székely. Pathology growth model based on particles. In *Proceedings of the 6th International Conference on Medical Image Computing and Computer Aided Intervention (MICCAI’2003)*, volume 2878 of *LNCS*, pages 25–32, 2003.
- [SBSG17] C. Savariaux, P. Badin, A. Samson, and S. Gerber. A comparative study of the precision of Carstens and Northern Digital Instruments electromagnetic articulographs. *Journal of Speech, Language and Hearing Research*, 60 :322–340, February 2017.

- [SD95] M. Stone and E. P. Davis. A head and transducer support system for making ultra-sound images of tongue/jaw movement. *Journal of Acoustical Society of America*, 98(6) :3107–3112, December 1995.
- [SDU07] G. Slabaugh, Q. Dinh, and G. Unal. A variational approach to the evolution of radial basis functions for image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2007)*, Minneapolis, (MN, USA), June 2007.
- [SFTP⁺94] D. Saint-Félix, Y. Troussset, C. Picard, C. Ponchut, R. Roméas, and A. Rougée. In vivo evaluation of a new system for 3-D computerized angiography. *Physics in Medicine and Biology*, 39 :583–595, 1994.
- [She99] A. Sherstyuk. Kernel functions in convolution surfaces : A comparative analysis. *The Visual Computer*, 15(4) :171–182, 1999.
- [SKO89] M. Schumacher, K. Kutluk, and D. Ott. Digital rotational angiography in neuroradiology. *American Journal of Neuroradiology*, 10 :644–649, 1989.
- [SMJC99] C. Shadle, M. Mohammad, P. Jackson, and J. Carter. Multi-planar dynamic magnetic resonance imaging : New tools for speech research. In *Proceedings of the 13th International Congress of Phonetic Sciences (ICPhS'1999)*, pages 623–626, 1999.
- [SMvW⁺09] M. Schaap, C. T. Metz, T. van Walsum, A. G. van der Giessen, A. C. Weustink, N. R. A. Mollet, C. Bauer, H. Bogunovic, C. Castro, X. Deng, E. Dikici, T. O'Donnell, M. Frenay, O. Friman, M. H. Hoyos, P. H. Kitslaar, K. Krissian, C. Kühnel, M. A. Luengo-Oroz, M. Orkisz, Ö. Smedby, M. Styner, A. Szymczak, H. Tek, C. Wang, S. K. Warfield, S. Zambal, Y. Zhang, G. P. Krestin, and W. J. Niessen. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. *Medical Image Analysis*, 13/5 :701–714, 2009.
- [SNB⁺08] C. Schumann, M. Neugebauer, R. Bade, et al. Implicit vessel surface reconstruction for visualization and CFD simulation. *International Journal of Computer Assisted Radiology and Surgery*, 2(5) :275–286, 2008.
- [SS02] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1) :7–42, 2002.
- [SSM⁺07] M. Schaap, I. Smal, C. Metz, T. Walsum, and W. Niessen. Bayesian tracking of elongated structures in 3D images. In Nico Karssemeijer and Boudewijn Lelieveldt, editors, *Information Processing in Medical Imaging*, volume 4584 of *Lecture Notes in Computer Science*, pages 74–85. Springer Berlin Heidelberg, 2007.
- [SSMS98] K. Satoh, J. Satomi, S. Matsubara, and Nagahiro S. Measurement of volume ratio to predict coil compaction, on aneurysmal embolization. *Interventional Neuroradiology*, 4(1) :179–182, 1998.
- [SSS⁺06] M. Schünke, E. Schulte, U. Schumacher, M. Voll, and K. Wesker. *Atlas d'anatomie prométhée : tête et neuro-anatomie*, volume 3. Pollina s.a. France, maloine edition, 2006.
- [STH⁺94] R.B. Schwartz, H.M. Tice, S.M. Hooten, L.H. Hsu, and P.E. Stieg. Evaluation of intracranial aneurysms with helical CT : Correlation with conventional angiography and MR angiography. *Radiology*, 192 :717–722, 1994.
- [STH96] B. H. Story, I. R. Titze, and E. A. Hoffman. Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America*, 100 :537–553, 1996.
- [Sto05] M. Stone. A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, 19(6-7) :455–501, September-November 2005.

- [SZ00] C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. *International Journal of Computer Vision*, 40(3) :199–233, 2000.
- [Tag99] H.D. Tagare. Shape-based nonrigid correspondence with application to heart motion analysis. *IEEE Transactions on Medical Imaging*, 18(7) :570–579, July 1999.
- [Tau91] G. Taubin. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13 :1115–1138, 1991.
- [TBC95] N. Tsingos, E. Bittar, and M.-P. Cani. Implicit surfaces for semi-automatic medical organ reconstruction. In *Computer Graphics International (CGI'1995)*, pages 3–15, 1995.
- [TBG⁺96] F. Turjman, K. Bendib, C. Girerd, J.C. Froment, and M. Amiel. Pre-therapeutic evaluation of intracranial aneurysms using three-dimensional angiography (3D morphometer) : Preliminary results. In W. Taki, L. Picard, and H. Kikuchi, editors, *Advances in interventional neuroradiology and intravascular neurosurgery*, pages 75–79. Elsevier Science, 1996.
- [TCM⁺96] R.K. Tu, W.A. Cohen, K.R. Maravilla, W.H. Bush, N.H. Patel, J. Eskridge, and H.R. Winn. Digital subtraction rotational angiography for aneurysms of the intracranial anterior circulation : Injection method and optimization. *American Journal of Neuroradiology*, 15 :1127–1136, 1996.
- [TdTF⁺07] J. A. Tyrrell, E. di Tomaso, D. Fuja, R. Tong, K. Kozak, R. K. Jain, and B. Roy-sam. Robust 3-D modeling of vasculature imagery using superellipsoids. *IEEE Transactions on Medical Imaging*, 26(2) :223–237, 2007.
- [TIA⁺02] S. Tamatani, Y. Ito, H. Abe, T. Koike, S. Takeuchi, and R. Tanaka. Evaluation of the stability of aneurysms after embolization using detachable coils : Correlation between stability of aneurysms and embolized volume of aneurysms. *American Journal of Neuroradiology*, 23(5) :762–767, 2002.
- [TK07] A. Tamrakar and B. Kimia. No grouping left behind : From edges to curve fragments. In *Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV'2007)*, 2007.
- [TKH⁺05] M. Teschner, S. Kimmerle, B. Heidelberger, G. Zachmann, L. Raghupathi, A. Fuhrman, M.-P. Cani, F. Faure, N. Magnenat-Thalmann, W. Strasser, and P. Volino. Collision detection for deformable objects. *Computer Graphics Forum*, 24 :61–81, 2005.
- [TMHF99] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment – a modern synthesis. In B. Triggs, A. Zisserman, and R. Szeliski, editors, *Proceedings of the International Workshop on Vision Algorithms (IWVA '1999)*, number 1883 in Lecture Notes in Computer Science, pages 298–372. Springer Berlin / Heidelberg, 1999.
- [Tos87] G. Toscani. *Systèmes de Calibration Optique et Perception du Mouvement en Vision Artificielle*. PhD thesis, Université de Paris Sud, 1987.
- [TYW01] A. Tsai, A. Yezzi, and A.S. Willsky. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Transactions on Image Processing*, 10(8) :1169–1186, August 2001.
- [TYWI⁺03] A. Tsai, A. Yezzi, W. Wells III, C. Tempany, D. Tucker, A. Fan, W.E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imimage using level sets. *IEEE Transactions on Medical Imaging*, 22(2) :137–154, February 2003.

- [UZV⁺10] M. Uecker, S. Zhang, D. Voit, A. Karaus, K.-D. Merboldt, and J. Frahm. Real-time MRI at a resolution of 20 ms. *NMR in Biomedicine*, 23(8) :986–994, 2010.
- [VPOS16] F. Vasconcelos, D. Peebles, S. Ourselin, and D. Stoyanov. Spatial calibration of a 2D/3D ultrasound using a tracked needle. *International Journal of Computer Assisted Radiology and Surgery*, 11 :1091–1099, 2016.
- [VRR⁺98] B.K. Velthuis, G.J.E. Rinkel, L.M.P. Ramos, T.D. Witkamp, J.W. Berkelbach van der Sprenkel, W.P. Vandertop, and M.S. van Leeuwen. Subarachnoid hemorrhage : Aneurysm detection and preoperative evaluation with CT angiography. *Radiology*, 208 :423– 430, 1998.
- [VSAG95] P.T. Vieco, W.P. Shuman, G.F. Alsofrom, and C.E. Gross. Detection of circle of Willis aneurysms in patients with acute subarachnoid hemorrhage : A comparison of CT angiography and digital subtraction angiography. *American Journal of Roentgenology*, 165 :425–430, 1995.
- [vWBN05] T. van Walsum, S.A. Baert, and W.J. Niessen. Guide wire reconstruction and visualization in 3DRA using monoplane fluoroscopic imaging. *IEEE Transactions on Medical Imaging*, 24(5) :612–623, May 2005.
- [VWI95] P. Viola and W.M. Wells III. Alignment by maximization of mutual information. In *Proceedings of the 5th International Conference on Computer Vision (ICCV’1995)*, pages 16–23, 1995.
- [WAC07] X. Wu, J. Allard, and S. Cotin. Real-time modeling of vascular flow for angiography simulation. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Proceedings of the 10th International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI’2007)*, volume 4791 of *Lecture Notes in Computer Science*, pages 557–565. Springer-Verlag, Berlin, Heidelberg, 2007.
- [WC07] W.C.K Wong and A.C.S Chung. Probabilistic vessel axis tracing and its application to vessel segmentation with stream surfaces and minimum cost paths. *Medical Image Analysis*, 11 :567–587, 2007.
- [WCL⁺98] Y. Wang, C. Chui, H. Lim, Y. Cai, and K. Mak. Real-time interactive simulator for percutaneous coronary revascularization procedures. *Computer Aided Surgery*, 3(5) :211–227, 1998.
- [WDBP⁺05] B. Wrobel-Dautcourt, M.-O. Berger, B. Potard, Y. Laprie, and S. Ouni. A low-cost stereovision based system for acquisition of visible articulatory data. In *Proceedings of the 5th Conference on Auditory-Visual Speech Processing (AVSP’05)*, July 2005.
- [Web12] S. Webb, editor. *Physics of Medical Imaging*. Series in Medical Physics and Biomedical Engineering. CRC Press, second edition, 2012.
- [Wei99] J. Weickert. Coherence-enhancing diffusion filtering. *International Journal of Computer Vision*, 31(2) :111–127, 1999.
- [WH94] A. P. Witkin and P. S. Heckbert. Using particles to sample and control implicit surfaces. In *Proceedings of SIGGRAPH 94*, pages 269–277, 1994.
- [WH00] A. Wrench and W. J. Hardcastle. A multichannel articulatory speech database and its application for automatic speech recognition. In *Proceedings of the 5th International Seminar on Speech Production (ISSP’2000)*, pages 305–308, 2000.
- [WHM13] J. Wu, Q. Hu, and X. Ma. Comparative study of surface modeling methods for vascular structures. *Computerized Medical Imaging and Graphics*, 37(1) :4–14, January 2013.
- [WHM⁺15] X. Wu, J. Housden, Y. Ma, B. Razavi, K. Rhode, and D. Rueckert. Fast catheter segmentation from echocardiographic sequences based on segmentation from corresponding X-ray fluoroscopy for cardiac catheterization interventions. *IEEE Transactions on Medical Imaging*, 34(4) :861–876, 2015.

- [WLK⁺11] X. Wu, V. Luboz, K. Krissian, S. Cotin, and S. Dawson. Segmentation and reconstruction of vascular structures for 3D real-time simulation. *Medical Image Analysis*, 15(1) :22–34, February 2011.
- [WNV00] O. Wink, W. J. Niessen, and M. A. Viergever. Fast delineation and visualization of vessels in 3D angiographic images. *IEEE Transactions on Medical Imaging*, 19(4) :337–346, 2000.
- [WPL⁺05] X. Wu, V. Pegoraro, V. Luboz, P. F. Neumann, R. Bardsley, S. Dawson, and S. Cotin. New approaches to computer-based interventional neuroradiology training. *Studies in Health Technology and Informatics*, 111 :602–607, 2005.
- [WR07] S. Worz and K. . Rohr. Segmentation and quantification of human vessels using a 3-D cylindrical intensity model. *IEEE Transactions on Image Processing*, 16(8) :1994–2004, August 2007.
- [WTO⁺05] D. H. Whalen, M. K. Tiede, D. J. Ostry, H. Lehnert-LeHouillier, E. Vatikiotis-Bateson, and D. S. Hailey. The Haskins optically corrected ultrasound system (HOCUS). *Journal of Speech, Language and Hearing Research*, 48 :543–553, June 2005.
- [WTW⁺01] P.M White, E.M. Teasdale, J.M. Wardlaw, , and V. Easton. Intracranial aneurysms : CT angiography and MR angiography for detection – prospective blinded comparison in a large patient cohort. *Radiology*, 219(3) :739–739, 2001.
- [WZWI04] S.K. Warfield, K.H. Zou, and W.M. Wells III. Simultaneous truth and performance level estimation (STAPLE) : An algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging*, 23(7) :903–921, 2004.
- [YCM⁺01] P. J. Yim, J. J. Cebral, R. Mullick, H. B. Marcos, and P. L. Choyke. Vessel surface reconstruction with tubular deformable model. *IEEE Transactions on Medical Imaging*, 20(12) :1411–1421, December 2001.
- [YDK99] N. Young, N.W.C. Dorsch, and R.J. Kingston. Pitfalls in the use of spiral CT for identification of intracranial aneurysms. *Neuroradiology*, 41 :93–99, 1999.
- [YJS06] A. Yilmaz, O. Javed, and M. Shah. Object tracking : A survey. *ACM Computing Surveys*, 38(4), 2006. Article No. 13.
- [Yur14] Ahmed Yureidini. *Robust blood vessel reconstruction for interactive medical simulations*. PhD thesis, Ecole doctorale Sciences pour l’Ingénieur (Lille), 2014.
- [ZG16] M. Zhang and P. Golland. Statistical shape analysis : From landmarks to diffeomorphisms. *Medical Image Analysis*, 33 :155–158, October 2016.
- [ZSA⁺03] X.Q. Zhang, H. Shirato, H. Aoyama, S. Ushikoshi, T. Nishioka, D.Z. Zhang, and K. Miyasaka. Clinical significance of 3D reconstruction of arteriovenous malformation using digital subtraction angiography and its modification with CT information in stereotactic radiosurgery. *International Journal of Radiation Oncology, Biology, Physics*, 57(5) :1392–1399, 2003.

Publications de l'auteur

Chapitres d'ouvrage

- [1] René Anxionnat, Marie-Odile Berger, and Erwan Kerrien. Time to go augmented in vascular interventional neuroradiology ? In Cristian Linte, Elvis Chen, Marie-Odile Berger, John Moore, and David Holmes III, editors, *Augmented Environments for Computer-Assisted Interventions*, volume 7815 of *Lecture Notes in Computer Sciences*, pages 3–8. Springer, 2013.
- [2] Yiyi Wei, Stéphane Cotin, Jérémie Dequidt, Christian Duriez, Jérémie Allard, and Erwan Kerrien. A (Near) Real-Time Simulation Method of Aneurysm Coil Embolization. In Yasuo Murai, editor, *Aneurysm*, pages 223–248. InTech, August 2012.
- [3] Michael Aron, Marie-Odile Berger, Erwan Kerrien, and Yves Laprie. Acquisition multimodale de données articulatoires. In Alain Cavé and Christian Marchal, editors, *L'imagerie médicale pour l'étude de la parole*, Traité Cognition et Traitement de l'Information, IC2, pages 175–196. Hermes Science Publications, November 2009.

Actes de conférences

- [4] Marie-Odile Berger, Erwan Kerrien, Gilles Simon, Antoine Tabbone, Laurent Wendling, and Brigitte Wrobel-Dautcourt, editors. *Actes des Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur - ORASIS 2003*. INRIA, 2003.

Revue internationale

- [5] Erwan Kerrien, Ahmed Yureidini, Jeremie Dequidt, Christian Duriez, René Anxionnat, and Stéphane Cotin. Blood vessel modeling for interactive simulation of interventional neuroradiology procedures. *Medical Image Analysis*, 35 :685 – 698, January 2017.
- [6] Michaël Aron, Marie-Odile Berger, Erwan Kerrien, Brigitte Wrobel-Dautcourt, Blaise Portard, and Yves Laprie. Multimodal acquisition of articulatory data : Geometrical and temporal registration. *Journal of the Acoustical Society of America*, 139(2) :638–648, 2016.
- [7] Nazim Haouchine, Stephane Cotin, Igor Peterlik, Jeremie Dequidt, Mario Sanz-Lopez, Erwan Kerrien, and Marie-Odile Berger. Impact of soft tissue heterogeneity on augmented reality for liver surgery. *IEEE Transactions on Visualization and Computer Graphics*, 21(5) :584–597, 2014.
- [8] Marie-Odile Berger, René Anxionnat, Erwan Kerrien, Luc Picard, and Michael Söderman. A methodology for validating a 3D imaging modality for brain AVM delineation : Application to 3DRA. *Computerized Medical Imaging and Graphics*, 32(7) :544–553, 2008.
- [9] Fleur Longuetaud, Jean-Michel Leban, Frédéric Mothe, Erwan Kerrien, and Marie-Odile Berger. Automatic detection of pith on CT images of spruce logs. *Computers and Electronics in Agriculture*, 44(2) :107–119, 2004.
- [10] René Anxionnat, Serge Bracard, Xavier Ducrocq, Yves Troussset, Laurent Launay, Erwan Kerrien, Marc Braun, Régis Vaillant, Francesco Scmazzone, Ariel Lebedinsky, and Luc

Picard. Intracranial aneurysms : Clinical value of 3D digital subtraction angiography in the therapeutic decision and endovascular treatment. *Radiology*, 218(3) :799–808, 2001.

- [11] L. Picard, Eric Maurincomme, M. Soderman, Jacques Feldmar, René Anxionnat, L. Lauenay, K. Ericson, Grégoire Malandain, S. Bracard, Erwan Kerrien, O. Flodmark, and Nicholas Ayache. X-ray angiography in stereotactic conditions : Techniques and interest for interventional neuroradiology. *Stereotactic and Functional Neurosurgery*, 68(1-4 Pt 1) :117–120, 1997.

Conférences internationales

- [12] Raffaella Trivisonne, Erwan Kerrien, and Stéphane Cotin. Augmented 3D catheter navigation using constrained shape from template. In *Proceedings of the 10th Hamlyn Symposium on Medical Robotics*, London, UK, June 2017.
- [13] Pierre-Jean Petitprez, Erwan Kerrien, and Pierre-Frédéric Villard. PoLAR : A portable library for augmented reality. In *Proceedings of the 15th IEEE International Symposium on Mixed and Augmented Reality (ISMAR'16)*, Merida, Mexico, septembre 2016.
- [14] Charlotte Delmas, Cyril Riddell, Yves Troussset, Erwan Kerrien, Marie-Odile Berger, René Anxionnat, and Serge Bracard. Intra-operative 3d micro-coil imaging using subsampled tomographic acquisition patterns on a biplane c-arm system. In *Proceedings of the 4th International Conference on Image Formation in X-Ray Computed Tomography (CT meeting'16)*, Bamberg, Germany, July 2016.
- [15] Charlotte Delmas, Marie-Odile Berger, Erwan Kerrien, Cyril Riddell, Yves Troussset, René Anxionnat, and Serge Bracard. Three-dimensional curvilinear device reconstruction from two fluoroscopic views. In *SPIE, Medical Imaging 2015 : Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9415, page 94150F, San Diego, CA, 2015.
- [16] Pierre-Frédéric Villard, Pierre Escamilla, Erwan Kerrien, Sébastien Gorges, Yves Troussset, and Marie-Odile Berger. Preliminary study of rib articulated model based on dynamic fluoroscopy images. In *SPIE Medical Imaging*, page 90361Y, February 2014.
- [17] Mario Sanz-Lopez, Jérémie Dequidt, Erwan Kerrien, Christian Duriez, Marie-Odile Berger, and Stéphane Cotin. Testbed for assessing the accuracy of interventional radiology simulations. In *Proceedings of the 6th International Symposium on Biomedical Simulation (ISBMS'2014)*, LNCS, Strasbourg, France, October 2014. Springer.
- [18] Nazim Haouchine, Jérémie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stéphane Cotin. Towards an accurate tracking of liver tumors for augmented reality in robotic assisted surgery. In *International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, June 2014.
- [19] Nazim Haouchine, Jérémie Dequidt, Igor Peterlik, Erwan Kerrien, Marie-Odile Berger, and Stéphane Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *ISMAR - IEEE International Symposium on Mixed and Augmented Reality 2013*, Adelaide, Australie, October 2013.
- [20] Ahmed Yureidini, Erwan Kerrien, Jérémie Dequidt, Christian Duriez, and Stéphane Cotin. Local implicit modeling of blood vessels for interactive simulation. In Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Morita, editors, *Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'2012)*, volume 7510 of *Lecture Notes in Computer Science*, pages 553–560, Nice, France, October 2012. Springer.
- [21] Ahmed Yureidini, Erwan Kerrien, and Stéphane Cotin. Robust RANSAC-based blood vessel segmentation. In Dave R. Haynor and Sébastien Ourselin, editors, *SPIE Medical Imaging*, volume 8314, page 8314M, San Diego, CA, United States, February 2012. SPIE Press.

- [22] Nazim Haouchine, Jérémie Dequidt, Erwan Kerrien, Marie-Odile Berger, and Stéphane Cotin. Physics-based augmented reality for 3D deformable object. In *VRIPHYS - Virtual Reality Interaction and Physical Simulation*, Darmstadt, Germany, December 2012.
- [23] Ting Peng, Erwan Kerrien, and Marie-Odile Berger. A shape base framework to segmentation of tongue contours from MRI data. In *35th IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP 2010*, pages 662 – 665, Dallas, United States, March 2010.
- [24] Jérémie Dequidt, Christian Duriez, Stéphane Cotin, and Erwan Kerrien. Towards interactive planning of coil embolization in brain aneurysms. In G.-Z. Yang, D.J. Hawkes, D. Rueckert, A. Noble, and C. Taylor, editors, *Medical Image Computing and Computer Assisted Intervention - MICCAI 2009*, volume 5761 of *Lecture Notes in Computer Science*, pages 377–385, London, United Kingdom, October 2009. Springer Berlin / Heidelberg.
- [25] Michael Aron, Asterios Toutios, Marie-Odile Berger, Erwan Kerrien, Brigitte Wrobel-Dautcourt, and Yves Laprie. Registration of multimodal data for estimating the parameters of an articulatory model. In *IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP 2009*, pages 4489 – 4492, Taipei, Taiwan, Province Of China, 2009. IEEE.
- [26] Jérémie Dequidt, Maud Marchal, Christian Duriez, Erwan Kerrien, and Stéphane Cotin. Interactive simulation of embolization coils : Modeling and experimental validation. In *Medical Imaging Computing and Computer Assisted Intervention -MICCAI'08*, volume 5241 of *Lecture Notes in Computer Science*, pages 695–702, New York City, United States, 2008. Springer.
- [27] Michael Aron, Anastasios Roussos, Marie-Odile Berger, Erwan Kerrien, and Petros Maragos. Multimodality Acquisition of Articulatory Data and Processing. In *16th European Signal Processing Conference - EUSIPCO 2008*, Lausanne, Switzerland, 2008. EPFL.
- [28] Michael Aron, Marie-Odile Berger, and Erwan Kerrien. Multimodal fusion of electromagnetic, ultrasound and MRI data for building an articulatory model. In *8th International Seminar On Speech Production - ISSP'08*, Strasbourg, France, 2008.
- [29] Michael Aron, Nicolas Ferveur, Erwan Kerrien, Marie-Odile Berger, and Yves Laprie. Acquisition and synchronization of multimodal articulatory data. In *8th Annual Conference of the International Speech Communication Association - Interspeech'07*, pages 1398–1401, Antwerpen, Belgium, 2007.
- [30] Sébastien Gorges, Erwan Kerrien, Marie-Odile Berger, Yves Troussset, and Jérémie Pescatore. An effective technique for calibrating the intrinsic parameters of a vascular C-arm from a planar target. In *SPIE Medical Imaging 2006*, pages 6141–69, San Diego, United States, February 2006.
- [31] Michael Aron, Erwan Kerrien, Marie-Odile Berger, and Yves Laprie. Coupling electromagnetic sensors and ultrasound images for tongue tracking : acquisition setup and preliminary results. In H. C. Yehia, D. Demolin, and R. Laboissiere, editors, *7th International Seminar on Speech Production - ISSP'06*, Ubatuba, Brazil, December 2006.
- [32] Erwan Kerrien, Jacques Hubert, Marie-Odile Berger, and Alain Blum. Registration and segmentation of multislice 3DCT images : Application to laparoscopic surgery of the upper urinary system. In H. U. Lemke, K. Inamura, and A.G. Farman, editors, *19th International Conference on Computer Assisted Radiology and Surgery - CARS'05*, volume 1281 of *International Congress Series*, page 1282, Berlin, Germany, June 2005. Elsevier.
- [33] Sébastien Gorges, Erwan Kerrien, Marie-Odile Berger, Yves Troussset, Jérémie Pescatore, René Anxionnat, and Luc Picard. Model of a vascular C-arm for 3D augmented fluoroscopy in interventional radiology. In James Duncan and Guido Gerig, editors, *8th International Conference on Medical Image Computing and Computer Assisted Intervention-MICCAI'05*, volume 3750 of *Lecture Notes in Computer Science*, pages 214–222, Palm Springs, United States, October 2005. Springer.

- [34] Sébastien Gorges, Erwan Kerrien, Marie-Odile Berger, Yves Troussel, Jérémie Pescatore, René Anxionnat, and Luc Picard. Calibration of the intrinsic parameters of a vascular C-arm for 3D enhanced fluoroscopy. In H.U. Lemke, K. Inamura, and A.G. Farman, editors, *19th International Congress and Exhibition Computer Assisted Radiology and Surgery - CARS'05*, International Congress Series, page 1307, Berlin, Germany, June 2005. Elsevier.
- [35] Marie-Odile Berger, René Anxionnat, and Erwan Kerrien. The use of 3D Xray angiographic images for volume determination of cerebral arteriovenous malformations. In *SPIE Medical Imaging 2004*, pages 634–643, San Diego, 2004.
- [36] Marie-Odile Berger, René Anxionnat, and Erwan Kerrien. A methodology for validating a new imaging modality with respect to a gold standard imagery : Example of the use of 3DRA and MRI for AVM delineation. In Chrisian Barillot, David R. Haynor, and Pierre Hellier, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2004*, volume 3216 of *Lecture Notes in Computer Science*, pages 516–524, Saint Malo, France, 2004. Springer-Verlag.
- [37] René Anxionnat, Marie-Odile Berger, Erwan Kerrien, Serge Bracard, and Luc Picard. Intra- and inter-observer variability in the angiographic delineation of brain arterio-venous malformations (AVMs). In *17th International Congress and Exhibition on Computer Assisted Radiology and Surgery - CARS'2003*, pages 1297–1298, Londres, Royaume-Uni, 2003. Elsevier.
- [38] Erwan Kerrien, Olivier Lévrier, René Anxionnat, Sylvain Sauty, and Jérôme Knoploch. Automated registration of 3D X-ray angiography images to magnetic resonance images. In K. Hanson, editor, *SPIE Medical Imaging 2001*, volume 4322/1, pages 623–632, San Diego, CA, USA, 2001. SPIE - The International Society for Optical Engineering, SPIE Press.
- [39] Erwan Kerrien, Laurent Launay, Marie-Odile Berger, Éric Maurincomme, Régis Vaillant, and Luc Picard. Fully automatic 3D/2D subtracted angiography registration. In H.U. Lemke, M.W. Vannier, K. Inamura, and A.G. Farman, editors, *13th International Congress and Exhibition Computer Assisted Radiology and Surgery - CARS'99*, International Congress Series, page 989, Paris, France, June 1999. Elsevier.
- [40] Erwan Kerrien, Marie-Odile Berger, Eric Maurincomme, Laurent Launay, Régis Vaillant, and Luc Picard. Fully automatic 3D/2D subtracted angiography registration. In Alan Colchester Chris Taylor, editor, *International Conference for Medical Image Computing & Computer Assisted Intervention - MICCAI'99*, volume 1679 of *Lecture Notes in Computer Science*, pages 664–671, Cambridge, England, 1999. Springer-Verlag.
- [41] Erwan Kerrien, Régis Vaillant, Laurent Launay, Marie-Odile Berger, and Eric Maurincomme. Machine precision assessment for 3D/2D digital subtracted angiography images registration. In *SPIE Medical Imaging'98*, pages 1–11, San Diego, USA, 1998.

Conférences nationales

- [42] Nazim Haouchine, Stéphane Cotin, Jérémie Dequidt, Erwan Kerrien, and Marie-Odile Berger. Réalité augmentée pour la chirurgie minimalement invasive du foie utilisant un modèle biomécanique guidé par l'image. In *Reconnaissance de Formes et Intelligence Artificielle (RFIA) 2014*, France, June 2014.
- [43] Ahmed Yureidini, Erwan Kerrien, and Stéphane Cotin. Reconstruction robuste des vaisseaux sanguins par surfaces implicites locales. In *Treizième congrès francophone des jeunes chercheurs en vision par ordinateur - ORASIS'11*, Praz-sur-Arly, France, June 2011.
- [44] Michael Aron, Erwan Kerrien, Marie-Odile Berger, and Yves Laprie. Fusion de capteurs électromagnétiques et d'échographies pour le suivi de la langue. In *Onzième congrès francophone des jeunes chercheurs en vision par ordinateur - ORASIS'07*, Obernai, France, June 2007.

- [45] Erwan Kerrien, Marie-Odile Berger, Jacques Hubert, and Alain Blum. Recalage et segmentation d'images scanographiques 3D : application à la représentation du haut appareil urinaire avant coeliochirurgie. In *Journées de recherche en imagerie médicale - JRIM'05*, page 26, Nancy, France, March 2005.
- [46] Erwan Kerrien, Marie-Odile Berger, Eric Maurincomme, Laurent Launay, Régis Vaillant, and Luc Picard. Recalage automatique d'images d'angiographie. In *Septième congrès francophone des jeunes chercheurs en vision par ordinateur - ORASIS'99*, page 9 p, Aussois, France, 1999.
- [47] Erwan Kerrien and Marie-Odile Berger. Étude de la précision de la machine pour le recalage 2D/3D d'images d'angiographie soustraite. In *Sixième congrès francophone des jeunes chercheurs en vision par ordinateur - ORASIS'97*, pages 39–48, octobre 1997.

Communications

- [48] Ahmed Yureidini, Jérémie Dequidt, Erwan Kerrien, Christian Duriez, and Stéphane Cotin. Computer-based simulation for the endovascular treatment of intracranial aneurysms. In *LIVIM Imaging Workshop*, Strasbourg, France, December 2011.
- [49] Michael Aron, Marie-Odile Berger, and Erwan Kerrien. Evaluation of the uncertainty of multimodal articulatory data. In *Ultrafest V*, New Heaven, United States, March 2010.
- [50] Erwan Kerrien, Marie-Odile Berger, and Jérémie Dequidt. Refining the 3D surface of blood vessels from a reduced set of 2D DSA images. In PJ "Eddie" Edwards, Kensaku Mori, and Tobias Sielhorst, editors, *AMI-ARCS 2008*, pages 61–69, New York, NY, United States, 2008.
- [51] Sébastien Gorges, Erwan Kerrien, Marie-Odile Berger, Yves Troussel, Jérémie Pescatore, René Anxionnat, Luc Picard, and Serge Bracard. 3D augmented fluoroscopy in interventional neuroradiology : Precision assessment and first evaluation on clinical cases. In *Workshop on Augmented environments for Medical Imaging and Computer-aided Surgery - AMI-ARCS 2006 (held in conjunction with MICCAI'06)*, Copenhagen, Denmark, November 2006. Wolfgang Birkfellner, Nassir Navab and Stephane Nicolau.

Conférences médicales

- [52] René Anxionnat, Federico Rocca, Serge Bracard, Jérémie Dequidt, Erwan Kerrien, Christian Duriez, Marie-Odile Berger, and Stéphane Cotin. Evaluation of a computer-based simulation for the endovascular treatment of intracranial aneurysms. In *10th congress of the World Federation of Interventional and Therapeutic Neuroradiology - WFITN 2009*, Montréal, Canada, June 2009.
- [53] Jacques Hubert, Alain Blum, Erwan Kerrien, and Marie-Odile Berger. Recalage et segmentation d'images scanographiques 3D : application à la représentation du haut appareil urinaire. In *99è Congrès Français d'Urologie*, Paris, France, November 2005. Association Française d'Urologie (AFU).
- [54] René Anxionnat, Erwan Kerrien, Marie-Odile Berger, Serge Bracard, and Luc Picard. Radiosurgical planning of brain arteriovenous malformations (AVMS) : A user interface for the delineation of 3D target from 2D angiographic projections. In *XXX Congress of the European Society of Neuroradiology*, volume 47 (suppl. 1), page S143, Barcelona, Spain, September 2005.
- [55] René Anxionnat, Jean-Claude Marchal, Alexandre Cordebar, Isabelle Buchheit, Erwan Kerrien, Yves Troussel, Marie-Odile Berger, Pierre Bey, Pierre Aletti, Serge Bracard, Sylvain Coulbois, Jean Auque, and Luc Picard. Radiosurgery of complex shaped brain arteriovenous malformations using multileaf collimator. Interest of 3D X-ray angiography and fusion with MRI for the delineation of the target. In *Symposium Neuroradiologicum*, volume 29/1, page 1S83, Paris, France, 2002.

- [56] René Anxionnat, Serge Bracard, Francisco Scomazzoni, Ariel Lebedinsky, Yves Troussset, Erwan Kerrien, and Luc Picard. Anévrismes et angiographie 3D. In *28^{ème} congrès de la Société Française de Neuroradiologie et 3^{ème} journée de Neuroradiologie de langue française (Bruxelles, Belgique)*, 2000. Conférence invitée.

Brevets

- [57] Sébastien Gorges, Yves Troussset, Jérémie Pescatore, Erwan Kerrien, and Marie-Odile Berger. Procédé de correction de recalage d'images radiographiques, 01 2009. Also filed in US (US2009022381 and US8180132) and Japan (JP2009022754).
- [58] Sébastien Gorges, Yves Troussset, Jérémie Pescatore, Vincent Bismuth, Marie-Odile Berger, and Erwan Kerrien. System and method to improve visibility of an object in an imaged subject, 10 2008. Also filed in US (US2008267490 and US7853061) and Japan (JP2008272470).
- [59] Sébastien Gorges, Yves Troussset, Jérémie Pescatore, Erwan Kerrien, Marie-Odile Berger, and Régis Vaillant. Acquisition geometry determining method for e.g. Roentgen system, involves computing set of calibration matrices in advance for computation of projection matrix, where each calibration matrix is assigned to calibration position of system, 06 2006. Also filed in Germany (DE102005059301), US (US2007172033 and US7559694) and Japan (JP2006167455).
- [60] Sébastien Gorges, Yves Troussset, Jérémie Pescatore, Erwan Kerrien, and Marie-Odile Berger. Acquisition geometry determining method for e.g. Roentgen system, involves computing set of calibration matrices in advance for computation of projection matrix, where each calibration matrix is assigned to calibration position of system, 08 2006. Also filed in Germany (DE102006007255) and Japan (JP2006223868).
- [61] Régis Vaillant, Dorothée Freymann, Laurent Launay, and Erwan Kerrien. Procédé et dispositif de reconstruction d'une image tridimensionnelle dynamique d'un objet parcouru par un produit de contraste, 02 2002. Also filed in Germany (DE10137521), US (US2002123680 and US7113631) and Japan (JP2002177254).
- [62] Yves Troussset, Laurent Launay, and Erwan Kerrien. Medical imaging system with a display giving a three-dimensional representation, uses processing of two-dimensional image data to produce a three-dimensional display with the same view point as the capture device, 10 2002. Also filed in US (US2003018250 and US7123255).
- [63] Erwan Kerrien, Sylvie Jacquot-Ingles, Éric Maurincomme, and Laurent Launay. Procédé pour le calibrage d'un système d'assistance pour intervention chirurgicale du type à imagerie médicale, 08 2002. Also filed in US (US2002122536).
- [64] Erwan Kerrien, Sylvie Jacquot-Ingles, Éric Maurincomme, and Laurent Launay. Procédé d'étalonnage d'un système d'imagerie, support de mémoire et dispositif associé, 10 2002. Also filed in Germany (DE10218760), Japan (JP2003019128) and US (US2002163996 and US6854884).
- [65] Éric Maurincomme and Erwan Kerrien. Method for matching three-dimensional radiological and nuclear magnetic resonance (3-D) images by matching an initial image point and then making two-dimensional projections of the 3-D images, 06 2001. Also filed in Germany (DE10059697), US (US2001036303 and US6879711) and Japan (JP2001212126).
- [66] Erwan Kerrien, Éric Maurincomme, Laurent Launay, and Régis Vaillant. Method for automatic registration of images, 03 2001. Also filed in Europe (EP1081647), US (US6714668) and Japan (JP2001101392).

Rapports

- [67] Y. Laprie, P. Maragos, A. Roussos, M. O. Berger, E. Kerrien, and M. Aron. Final report on design, acquisition and processing of articulatory data. Technical report, FET ASPI Report (EU Project no. 2005-021324), 2009.
- [68] M. Aron, M. O. Berger, and E. Kerrien. Initial report on the design of the acquisition system. Technical report, FET ASPI Report (EU Project no. 2005-021324), 2006.

Logiciels

- [69] Erwan Kerrien, Pierre-Frédéric Villard, and Pierre-Jean Petitprez. PoLAR : a portable Library for Augmented Reality. Distribué sous licence GPL, <http://polar.inria.fr>, 2015.
- [70] E. Kerrien, G. Simon, M-O. Berger, and V. Lepetit. RALIB : une bibliothèque logicielle pour le traitement d’images, l’imagerie médicale, et la réalité augmentée. Dépôt à l’Agence pour la Protection des Programmes, numéro IDN.FR.001.100005.000.R.P.2004.000.10000, 2004.

Thèses et mémoires

- [71] Erwan Kerrien. *Outils d’imagerie multimodalités pour la neuroradiologie interventionnelle*. PhD thesis, Institut National Polytechnique de Lorraine, 2000.
- [72] Erwan Kerrien. Tomographie d’impédance électrique appliquée à l’imagerie fonctionnelle de l’estomac. Rapport de mémoire ingénieur, École Nationale Supérieure des Télécommunications de Paris, 1994. In French.
- [73] Erwan Kerrien. Inversion de données géophysiques par la méthode des projections sur les ensembles convexes / Geophysical data inversion using projection onto convex sets. Rapport de stage ingénieur, École Nationale des Télécommunications de Paris, 1994. In English.

Vulgarisation scientifique

- [74] Sylvie Alayrangues, Gilles Dowek, Erwan Kerrien, Jean Mairesse, and Thierry Viéville. Médiation en sciences du numériques : un levier pour comprendre notre quotidien ? In *Science & You*, Nancy, France, June 2015.
- [75] Antoine Rousseau, Aurélie Darnaud, Brice Goglin, Céline Acharian, Christine Leininger, Christophe Godin, Clarisse Holik, Claude Kirchner, Diane Rives, Elodie Darquie, Erwan Kerrien, Fabrice Neyret, Florent Massegia, Florian Dufour, Gérard Berry, Gilles Dowek, Hélène Robak, Hélène Xypas, Irina Illina, Isabelle Gnaedig, Joanna Jongwane, Jocelyne Ehrel, Laurent Viennot, Laure Guion, Lisette Calderan, Lola Kovacic, Marie Collin, Marie-Agnès Enard, Marie-Hélène Comte, Martin Quinson, Martine Olivi, Mathieu Giraud, Mathilde Dorémus, Mia Ogouchi, Muriel Droin, Nathalie Lacaux, Nicolas Rougier, Nicolas Roussel, Pascal Guitton, Pierre Peterlongo, Rose-Marie Cornus, Simon Vandermeersch, Sophie Maheo, Sylvain Lefebvre, Sylvie Boldo, Thierry Viéville, Véronique Poiriel, Aline Chabreuil, Arnaud Fischer, Claude Farge, Claude Vadel, Isabelle Astic, Jean-Pierre Dumont, Loïc Féjoz, Patrick Rambert, Pierre Paradinas, Sophie De Quatrebarbes, and Stéphane Laurent. Médiation Scientifique : une facette de nos métiers de la recherche. Rapport interne Inria, March 2013.
- [76] Stéphane Cotin and Erwan Kerrien. La simulation en radiologie interventionnelle. *DocSciences*, 13 :10–15, October 2010.

Résumé. Les travaux de recherche repris de manière synthétique dans ce mémoire, abordent les questions liées à l'imagerie médicale multimodale et interventionnelle dans deux contextes applicatifs. Le premier traite de l'acquisition de données multimodales dynamiques (son, images échographiques, capteurs de mouvement électromagnétiques, ...) chez un locuteur. Des procédés de recalage, de synchronisation et de segmentation de ces données sont décrits pour bâtir des modèles du conduit vocal associés à une variabilité statistique. Le deuxième s'intéresse à la neuroradiologie interventionnelle, thérapie guidée par l'image radiologique. En vue d'améliorer l'expérience visuelle du praticien, un premier ensemble de travaux traite de l'intégration de l'angiographie rotationnelle 3D dans la salle d'opération d'une part en la superposant aux images per-opératoires, et d'autre part en y remplaçant le micro-outil chirurgical reconstruit en temps interactif à partir de deux vues radiologiques. Enfin, un modèle d'arbre vasculaire 3D est présenté, adapté à une simulation physique interactive du geste thérapeutique. La voie est ainsi ouverte au développement de modèles plus réalistes et plus prédictifs de la complexité de la scène chirurgicale. L'objectif est à terme d'offrir une vue 3D et dynamique de cette scène pendant l'opération.

Abstract. The research works briefly reviewed in this dissertation deal with questions related to multimodal and interventional medical imaging within two application frameworks. First, multimodal dynamic data acquisition of a speaking person is addressed (sound, ultrasound images, electromagnetic motion capture sensors,...). Registration, synchronization and segmentation techniques are described in a view to build statistically founded models of the vocal tract. A second context is provided by interventional neuroradiology, an image-guided therapy technique. The goal is to improve the interventionalist's visual experience through a better integration of the 3D rotational angiography into the operating room. It is first shown how this 3D pre-operative image can be superimposed onto the live intra-operative video image feed. Besides, a real-time algorithm is developed to reconstruct the surgical micro-tool in 3D from two X-ray views, to locate it in the 3D image. Then, a new 3D model of the vessel tree is presented, that is adapted to interactive simulation of the therapeutic gesture. This paves the way towards more realistic and predictive models of the surgical scene. The long term objective is to make a 3D and dynamic view of this scene available to the interventionalist during the operation.